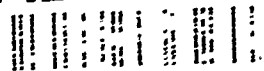


AD-A248 771



UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1991

GRADUATE STUDENT RESEARCH PROGRAM
(GSRP) REPORTS

VOLUME 8

ROME LABORATORY
ARNOLD ENGINEERING DEVELOPMENT CENTER
F. J. SEILER RESEARCH LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 UPLANDER WAY
CULVER CITY, CA 90230-6608

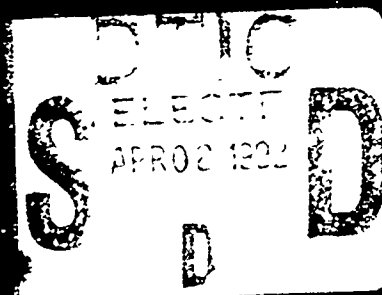
SUBMITTED TO:

LT. COL. CLAUDE CAVENDER
PROGRAM MANAGER

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

BOLLING AIR FORCE BASE
WASHINGTON, D.C.

DECEMBER 1991



AEOSR-TR- 92 0175

AIR FORCE OF SCIENTIFIC RESEARCH (AFSC)
REPORT SUBMITTED TO DTIC
This report has been reviewed and is
approved for public release IAW AFR 190-12
distribution unlimited.
Signed Program Manager

Approved for public release.
distribution unlimited

92 4 08 012

92-09033



REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, completing and reviewing the collection of information, sending comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

| | | | | | |
|---|--|---|--|--|--|
| 1. AGENCY USE ONLY (Leave blank) | | 2. REPORT DATE 9 January 1992 | | 3. REPORT TYPE AND DATES COVERED 30 Sep 90-30 Sep 91 | |
| 4. TITLE AND SUBTITLE 1991 Graduate Student Research Program (GSRP) Volumes 6-9 Vol. 8 | | | | 5. FUNDING NUMBERS F49620-90-C-0076 | |
| 6. AUTHOR(S) Mr Gary Moore | | | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research Development Laboratories (RDL) 5800 Uplander Way Culver City CA 90230-6608 AFOSR-TR | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER y2 0175 | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NI Bldg 410 Bolling AFB DC 20332-6448 Lt Col V. Claude Cavender | | | | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER | |
| 11. SUPPLEMENTARY NOTES | | | | | |
| 12a. DISTRIBUTION / AVAILABILITY STATEMENT UNLIMITED | | | | 12b. DISTRIBUTION CODE | |
| 13. ABSTRACT (Maximum 200 words) This program was started in 1982 as an adjunct to the SFRP. Its objectives are to permit graduate students to participate in research under the direction of a faculty member at an Air Force laboratory; stimulate professional association among graduate students, their supervising professors, and professional peers in the Air Force; to further research objectives of the Air Force; and to expose graduate students to potential thesis topics in areas of interest to the Air Force. During the summer of 1991 142 graduate students performed research for 10 weeks at Air Force laboratories. Their reports were submitted to RDL and consolidated into this annual report. | | | | | |
| 14. SUBJECT TERMS | | | | 15. NUMBER OF PAGES | |
| | | | | 16. PRICE CODE | |
| 17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED | | 18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED | | 19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED | |
| | | | | 20. LIMITATION OF ABSTRACT UL | |

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1991
GRADUATE STUDENT RESEARCH PROGRAM (GSRP) REPORTS

VOLUME 8

ROME LABORATORY
ARNOLD ENGINEERING DEVELOPMENT CENTER
F. J. SEILER RESEARCH LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Lt. Col. Claude Cavender

Program Manager, RDL
Claude Baum

Program Administrator, RDL
Gwendolyn Smith

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1991

| | |
|---------------------|----------------------|
| Accession For | |
| NTIS | CRA&I |
| DTIC | TAS |
| Unannounced | |
| Justification | |
| By | |
| Distribution/ | |
| Availability Code | |
| Dist | Avail and/or Special |
| A-1 | |



PREFACE

Reports in this document are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

This document is one of a set of 13 volumes describing the 1991 AFOSR Summer Research Program. The following volumes comprise the set:

| <u>VOLUME</u> | <u>TITLE</u> |
|----------------------|---|
| 1 | Program Management Report |
| | <i>Summer Faculty Research Program (SFRP) Reports</i> |
| 2 | Armstrong Laboratory, Wilford Hall Medical Center |
| 3 | Phillips Laboratory, Civil Engineering Laboratory |
| 4 | Rome Laboratory, Arnold Engineering Development Center, Frank J. Seiler Research Laboratory |
| 5 | Wright Laboratory |
| | <i>Graduate Student Research Program (GSRP) Reports</i> |
| 6 | Armstrong Laboratory, Wilford Hall Medical Center |
| 7 | Phillips Laboratory, Civil Engineering Laboratory |
| 8 | Rome Laboratory, Arnold Engineering Development Center, Frank J. Seiler Research Laboratory |
| 9 | Wright Laboratory |
| | <i>High School Apprenticeship Program (HSAP) Reports</i> |
| 10 | Armstrong Laboratory |
| 11 | Phillips Laboratory, Civil Engineering Laboratory |
| 12 | Rome Laboratory, Arnold Engineering Development Center |
| 13 | Wright Laboratory |

1991 GRADUATE STUDENT RESEARCH REPORTS

Rome Laboratory, Arnold Engineering Development Center, Frank J. Seiler Research Laboratory

| <u>Report Number</u> | <u>Report Title</u> | <u>Author</u> |
|---|---|-------------------|
| <u><i>Rome Laboratory</i></u> | | |
| Rome Air Development Center -- Griffiss (RADC) | | |
| 1 | The Role of Cut-Vertex Set Analysis in the Integrated Communications Network Management System (IMS) | Benjamin Hoe |
| 2 | Demonstration of a Low Energy Electron Diffraction System | Yolanda Kime |
| 3 | Analysis of the Electromigration-Induced Failure in the VLSI Interconnection Components and the Multisection Interconnections | Matthew Leipnitz |
| 4 | Windowing Comparison Project: The Effect of Window Shape and Size on Phoneme Identifiability | Beth Losiewicz |
| 5 | RTS Prototyper's Workbench: A Tool for Rapid Prototyping Provably Correct Real-Time Systems | James Peters, III |
| 6 | Experimental Evaluation of Optical Switching Technologies | Dean Richardson |
| 7 | Clock Synchronization Techniques: A Survey | Waleed Smari |
| 8 | The Characterization of Radar Clutter as an SIRP | Charles Widener |
| Rome Air Development Center -- Hanscom (RADCH) | | |
| 9 | Direct Excitation on Microwave-Spin Dressed States using a Laser-Excited Resonance Raman Interaction | Kuang Cheng |
| 10 | Millimeter-Wave Noise Modeling Investigation | Steven Lardizabal |
| 11 | Not Available at this Time | Thomas Lusby |
| 12 | Testings on Effects of Thermal Treatment on GaAs (and InP) Substrates at MOCVD Temperature | Yonghuan Zhou |
| <u><i>Arnold Engineering Development Center</i></u> | | |
| 13 | Non-Intrusive Testing of Composite Aircraft Engine Components: II | James Abbey |
| 14 | Parallel Signal Processing for Turbine Engine Testing | Ben Abbott |
| 15 | Wake and Projectile Velocity Estimation | George Aboutanos |
| 16 | An Extended Kalman Filter Observer for an Altitude Test Cell | W. Brian Ball |

**Report
Number**

Report Title

Author

Arnold Engineering Development Center (cont.)

- | | | |
|----|--|---------------------|
| 17 | Parallel Signal Processing for Turbine Engine Testing | Theodore Bapty |
| 18 | Development of a Computational Model of the Laser-Induced Fluorescence (LIF) of Hydrogen (H ₂) with Excitation at 193 Nanometers | Clinton Benefield |
| 19 | PLIF Development for Wake Physics Analysis | John Dempsey, Jr. |
| 20 | Parametric Evaluation of Finite Element Modeling Techniques | Dollena Hawkins |
| 21 | The Effect of Doppler Shift on the Apparent Plume Infrared Signatures of Gas Turbine Engines | Charles Hewitt, Jr. |
| 22 | Recommendations Report on Water Tunnel Analysis of ASTF C1 Plenum Section | Stephen Howard |
| 23 | Calibration Technique for COCODEC | Emily Joy |
| 24 | Dynamic Modeling of a Dual Flow Path Compression System | Jules Lindau V |
| 25 | Multigraph Implementation of Image Morphology | Michael Moore |
| 26 | An Approximate Method for the Prediction of Underexpanded Rectangular Nozzle Exhaust Plume Boundaries | Kyle Nash |
| 27 | A Local Lagrangian Finite Volume Model for the Infinite Domain Shock Tube Problem | Blair Rollin |
| 28 | Computer Model for Cavity Oscillations | Daniel Schatt |
| 29 | Monitoring an Axial Flow Compressor for Rotating Stall and Surge | John Sebghati |
| 30 | Demonstration of New Finite-Rate Chemistry CFD Code | Paul Vitt |
| 31 | Boundary Element Acoustic Analysis of the Exhaust Gas Management System in the AEDC Aeropropulsion Systems Test Facility | Michael Weaver |
| 32 | The Influence of Material Properties and Shape on the Measured Heat-Flux Distribution Using a Hemispherical Coaxial Thermal Couple Temperature Probe | William Wilk |

Frank J. Seller Research Laboratory

- | | | |
|----|--|-------------------|
| 33 | The Effects of Passive Porous Surface on Flow Field Development Under both Static and Dynamic Conditions | Gregory Addington |
| 34 | Active/Passive Control Synergism for a Planar Two-Dimensional Truss | Jeffrey Curtis |
| 35 | Optimal Placement of Passive Dampers Via Simulated Annealing | Tami Hamernik |

**Report
Number**

Report Title

Author

Frank J. Seiler Research Laboratory (cont.)

- | | | |
|----|---|-----------------------|
| 36 | Second Harmonic Generation in Poled Amorphous Quartz | Karl Kauffmann |
| 37 | Three-Dimensional Flow Structure Development and Effect Near the Root of an Oscillating Wing | John Klinge |
| 38 | An Investigation of Actively Controlled Static and Dynamic Airfoils | Julie Lovato |
| 39 | Relaxation Studies of Molten Salts | Maureen Parrish |
| 40 | Downwash Flow Mechanisms on a Pitching Canard/Wing Configuration | Dennis Strickland Jr. |
| 41 | An Ab Initio Study of the Adducts of Aluminum Hydrides, Halides, Hydroxides, and Oxides with HF and HCL | Marty Wilson |

1991 USAF SUMMER RESEARCH PROGRAM

sponsored by the

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by the

Rome Laboratory

FINAL REPORT

THE ROLE OF CUT-VERTEX SET ANALYSIS IN
THE INTEGRATED COMMUNICATIONS NETWORK MANAGEMENT SYSTEM (IMS)

| | |
|------------------|--------------------------------------|
| Prepared by: | Benjamin W. Hoe |
| Department and | Department of Electrical Engineering |
| University: | Polytechnic University, New York |
| Research Site: | Rome Laboratory/C3DA |
| | Griffiss AFB, NY 13441-5700 |
| USAF Researcher: | Nick Kowalchuk, Project Engineer |
| Date: | August 2, 1991 |
| Contract No: | F49620-90-C-0076 |

THE ROLE OF CUT-VERTEX SET ANALYSIS IN
THE INTEGRATED COMMUNICATIONS NETWORK MANAGEMENT SYSTEM (IMS)

by

Benjamin W. Hoe

ABSTRACT

Rome Laboratory is currently developing the Integrated Communications Network Management System (IMS) prototype, and cut-vertex set analysis plays an important role in the IMS prototype development process. This paper discusses the role of cut-vertex sets in terms of network survivability and reliability. A computer program for generating cut-vertex sets was developed, and comprised the major effort of this research. It also presents the theory behind the operation of the cut-vertex set program, how the program can be applied to the analysis of network survivability and reliability and it concludes with a discussion of possible applications to the IMS prototype design.

Acknowledgements

I would like to thank the Air Force Office of Scientific Research for sponsoring this research. In particular, I would like to thank Mr. Kris Newport and Dr. Mike Schroeder of the Mitre Corporation at Rome Lab for their invaluable technical information which was extremely helpful in this project.

I am also indebted to Nick Kowalchuk, the project engineer of IMS development, for the time he spent on our countless lengthy discussions through this project. His advice, guidance, comments and encouragement were invaluable. Mr. Joseph Zdanowicz must be mentioned for providing me with all administrative assistance during my research period in Rome Lab. In addition, I am very grateful to Dar. Hague, Brian Hendrickson, Priscilla Cassidy, Dr. Wayne Smith of Mississippi State University and other staff members of Network Design Lab for providing facility needed for this project, their assistance and helps.

I. Introduction:

The IMS is an "intelligent" communications network management system which is responsible for the optimal allocation of user service requests (voice, data, video) to available communication resources, so that in a wartime scenario users are guaranteed the highest possible level of services satisfied. The IMS is a distributed management system, with independent managers exchanging management information so that system level parameters can be optimized. The goal is for the IMS to be a survivable system that is able to adopt to a rapidly changing environment. Cut-vertex set analysis can provide vital information on vulnerabilities in the connectivity of a network, which can be used by the IMS to perform management actions that correct these vulnerabilities, and thus maintain a high degree of survivability.

A cut-vertex set is a set of nodes, which if removed from a network, will result in partial or complete isolation of some part of network from other parts (see Figure(2) and Table(1) on page(6)). It is important to compute the cut-vertex set so that the measures of network survivability such as Node Connectivity Factor (NCF), Link Connectivity Factor (LCF), Node Decomposition (ND) index and Link Tree (LT) index, and reliability factors such as Mean Time Between Failures (MTBF), Mean Up Time (MUT) and Mean Down Time (MDT) can be quantified. The Mitre Corporation at Rome Laboratory has done extensive work [4], [5] and [6] on computation of NCF, LCF, LT

and ND. Some of their work is briefly reviewed in this paper because of its critical role in the IMS.

II. Objective of The Research:

The survivability of network nodes is affected by their connectivity. During an attack, the enemy will be trying to identify and destroy the weak points of our Command, Control, Communications and Intelligence (C³I) system which means survivable network must not have single points of failure which would isolate the remainder of the network assets if it was destroyed. Such critical points must be identified and the network topology should be reconfigured so the network has higher degree of survivability. The ideal is to construct a topology which is survivable and has a favorable distribution of traffic load among its links. The IMS is a distributed management system; each manager in different domains of the IMS needs the updated knowledge of resource availability, hence, an optimal mapping of services can be provided to the users even when communication assets are degraded or lost [4]. The degree of the connectivity of a network is quantified by NCF, LCF, LT and ND. The LCF and NCF quantify the global measures in terms of the average number of links (or) nodes required to be removed in order to force the network into a stand alone configuration. The LCF and NCF can be obtained from following formulas:

$$LCF = \sum_{i=1}^n LCF_i \left(\frac{S_i}{S} \right) = \sum_{i=1}^n \frac{T_i (N_i - 1)}{E_i} \cdot \frac{(S_i)}{S} \text{-----} (1)$$

$$NCF = \sum_{i=1}^n NCF_i = \sum_{i=1}^n \sum_{j=1}^{M_i} N_j P(N_j) \text{-----} (2)$$

where

LCF_i : Link connectivity factor for component i

S_i : Number of edges in spanning tree for component i

S : Number of edges in spanning tree for the network

T_i : Number of spanning trees in component i

E_i : Number of edges in component i

N_i : Number of nodes in component i

n : Number of components

NCF_i : Node connectivity factor for component i

N_j : Number of nodes removed for the jth decomposition
path

$P(N_j)$: Probability of the jth decomposition path

M_i : Total paths formed in the component's
decomposition diagram

The computation of LT and ND is relatively trivial once LCF and NCF are obtained. The detailed computation of LT and ND are given in Newport and Schroeder [6].

In this research, the concentration was in obtaining the minimum cut-vertex set since it is the first to be considered

in the process of quantifying the survivability and reliability of an IMS managed system. The computation of cut-vertex sets is a time consuming process. This computational complexity is due to the size of the minimum cut-vertex search space and the decomposition factor (the average number of minimum cut-vertex sets actually found at each level of decomposition). Consider a 20 node network, the size of the cut-vertex search space is calculated using simple probability theory.

The total number of all possible node set (combination) is:

$$20 + \binom{20}{2} + \binom{20}{3} + \dots + \binom{20}{19} + 1 \text{-----} (3)$$

and

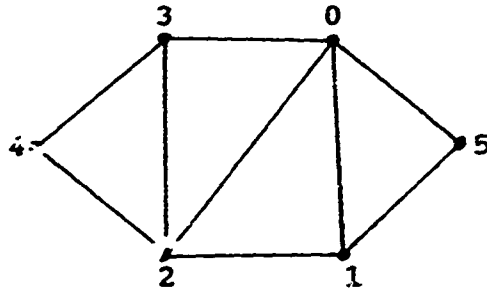
$$\binom{n}{k} = \frac{n!}{k! (n-k)!} \text{-----} (4)$$

Hence, the total number of all possible combination sequences will be $20 + 190 + 1140 + 4845 + \dots + 20 + 1$ which approaches infinity as the size of the network increases. The Mitre Corporation has developed a FORTRAN computer program to calculate all survivability parameters under Air Force contract No. F19628-86-C-0001. Other new ideals and methods such as knowledge-based approach and hybrid computations also have been published in recent Proceeding of the 1990 IEEE Military Communications Conference [7][8].

III. Details of The Research Effort:

In this section, the technique used in the computer program for generating minimum cut-vertex sets is discussed in detail and the flow charts are provided in the appendix. Then, the use of cut-vertex set analysis in quantifying survivability and reliability factors is discussed. The computer program was developed in "C" using the following steps:

Step (1): The user enters the network parameters such as total number of nodes, total number of links and individual links in terms of nodes (eg. link=(node1, node2)). Then, decomposition is performed according to data pattern sequences (to be discussed later), and the temp array is labeled for every node inside the node set being considered. Three different labels are used in the labelling process: a '-1' label is used if the node being considered is in array edv1, a '-2' label is used if the node being considered is in array edv2 and a '0' label is used for the situation when the nodes in both arrays edv1 and edv2 are the same as the nodes in the node set being considered. For example, Let's consider a node set containing node 0 and 1 such as {0,1}. For the given network in Figure (1), the temp array will be labeled as in (6). All links are stored by joining corresponding elements in the array end vertex 1 (edv1) and the end vertex 2 (edv2) (eg. link1 = (edv1(1),edv2(1))).



Figure(1)

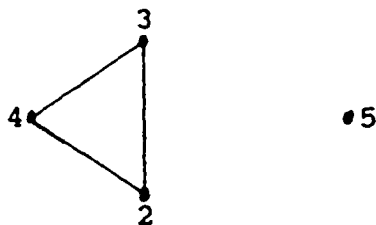
node set being removed = {0, 1}

$$\text{edv1}[i] = [0 \ 2 \ 1 \ 0 \ 2 \ 2 \ 3 \ 0 \ 1] \quad (5.1)$$

$$\text{edv2}[i] = [1 \ 0 \ 2 \ 3 \ 3 \ 4 \ 4 \ 5 \ 5] \quad (5.2)$$

$$\text{temp}[i] = [0 \ -2 \ -1 \ -1 \ . \ . \ . \ -1 \ -1] \quad (6)$$

The array temp[1] has a '0' label because both edv1[1] and edv2[1] have nodes that are elements of the node set {0,1}. The temp[2] and temp[3] array have label '-2' and '-1' respectively because node '0' is in edv2[2] and node '1' is in edv1[3]. By inspecting the array temp, one can tell which part of the network is affected by removing of the node set being considered.



Figure(2)

| <u>i</u> | <u>Cutset (C_i)</u> | <u>N_i</u> |
|----------|-------------------------------|----------------------|
| 1 | 0 1 | 2 |
| 2 | 0 2 | 2 |
| 3 | 2 3 | 2 |
| 4 | 0 1 2 | 3 |
| 5 | 0 1 3 | 3 |
| 6 | 0 1 4 | 3 |
| 7 | 0 2 3 | 3 |
| 8 | 0 2 4 | 3 |
| 9 | 1 2 3 | 3 |
| 10 | 0 1 2 3 | 4 |

Table (1)

Step (2): After the labelling of the temp array, some nodes have remained intact and these nodes are used to form subgraphs. In the above example, subgraph with nodes {2, 3, 4} and subgraph with node {5} are formed as in Figure(2) due to the fact these nodes were not part of the node set that are removed.

Step (3): In this final step, the program simply inspects whether all the subgraphs are connected or not. If some subgraph (or subgraphs) are not connected (stand alone status) then the node set being decomposed is a cut-vertex set. Otherwise, if all subgraphs are remain connected, the program will move on to consider another node set. Table(1) shows all cut-vertex sets for the network in Figure(1). The minimum cut-vertex sets for this network are C_1 , C_2 and C_3 as shown in Table (1).

The above process from step (1) through (3) is repeated until all minimum cut-vertex sets are obtained. All cut-vertex sets (not only minimum cut-vertex sets) can also be found by using the same procedure (1) through (3) with simple modification. The complete 'C' code for this program can be obtained as a Technical Memo in technical library of the Rome Laboratory at Griffiss Air Force Base.

In step (1), decomposition is performed according to a data sequence pattern from a data file. This data pattern is pre-generated and has the form shown below.

| | | | | | | | | |
|----|-----|-----|-----|---|---|---|---|-----|
| 0, | 1, | 2, | 3, | . | . | . | . | n |
| 0 | 01, | 02, | 03, | . | . | . | . | 0n, |
| 0 | 0 | 12, | 13, | . | . | . | . | 1n, |
| . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . |

The data file approach is used instead of direct generation of the node set in the main program for decomposition because of its simplicity and computational efficiency. The decomposition will be performed as in the given matrix; the program considers one node at a time and if no cut-set is obtained, then two nodes at a time will be considered, and thus the program will continue to increase the size of the node sets until all cut-vertex sets are obtained, or there are no more nodes to consider.

The computational time of finding cut-vertex sets is increased as the size of the network increases. It is also increased as the number of nodes in a cut-vertex set increases. However, large size of cut-vertex set is not critical in the IMS. As we can see in the NCF formula in (2), NCF is directly proportional to N_j , the number of nodes removed for the j th decomposition path, which is proportional to the number of nodes in the j th cut-vertex set. Therefore, for a cut-vertex set with a large number of nodes, the NCF is also large which implies that the number of nodes to be removed in order to force the network into stand alone

configuration is also large. Consequently, the network is survivable and a cut-vertex set with large number of nodes can be ignored.

So far the discussion has been limited to survivability parameters. The robustness of a network can also be understood by looking at its reliability parameters such as Mean Time between Failures(MTBF), Mean Up Time (MUT) and Mean Down Time (MDT). The computation of these reliability parameters is simple once the cut-vertex sets are found. In the papers [2] and [11], cut-link sets are used in the calculation of MTBF, MUT, MDT, however, cut-vertex sets are used to calculate these reliability parameters in this research since this information is already available from the previous effort of calculating NCF. Furthermore, communication nodes are more critical than individual links in a military communications system (as enemy is more likely to target a switching center rather than on a single communication line).

From the definition of a cut-vertex set, if all nodes in the cut-vertex set have failed, then the network is said to be disconnected into at least two subgraphs (or) we can define it as failure of the network in terms of full connectivity. The event of the failure of a network is defined as in following equation--

$$F = E_1 + E_2 + E_3 + \dots + E_n \quad (7)$$

where F is the event of the failure of a network and E_1 through E_n are the events in cut-vertex set 1 through cut-vertex set n . It follows that the probability of the failure of a network is:

$$\begin{aligned} P(F) &= P(E_1) + P(E_2) + \dots + P(E_n) \\ &= \sum_i P(E_i) - \sum_{i,j} P(E_i \cap E_j) + \sum_{i,j,k} P(E_i \cap E_j \cap E_k) \dots \end{aligned} \quad (8)$$

By using the probability theory, the upper bound of the $P(F)$ is defined as:

$$\begin{aligned} P(F) &\leq UB \\ \text{where } UB &= \sum_{i=1}^M p^{N_i} \end{aligned} \quad (9)$$

$$\begin{aligned} \text{and the lower bound is: } P(F) &\geq LB \\ \text{where } LB &= \sum_{i=2}^M \sum_{j=1}^{i-1} p^{N_i + N_j - L_{ij}} \end{aligned} \quad (10)$$

$$\text{and } p = \mu_d / (\mu_u + \mu_d)$$

from (9) and (10), the expression for $P(F)$ can be written as:

$$LB \leq P(F) \leq UB \quad (11)$$

In the above formulas N_i represents the number of elements (nodes) in the i th cut-vertex set, M represents the number of cut-vertex sets, E_i represents the events inside the i th cut-vertex set and L_{ij} represents the number of nodes in the intersection set of the i th cut-vertex set and the j th

cut-vertex set. As the above formulas indicate, upper bound and lower bound will be very small for a large cut-vertex set since probability is always less than 1 in (9) and (10). Therefore, large cut-vertex sets have little effect on the reliability of a network as it did in survivability. The failure rate of the i th cut-vertex set can be defined as

$$R_i = \frac{N_i}{\mu_u + \mu_d} p^{N_i - 1} \text{-----} (12)$$

where N_i and p are as defined above, and μ_u is the mean time spent in operational state, μ_d is the mean time spent in failure (detailed computation of μ_u and μ_d can be found in Plotkin and Einhom [11]). For a reasonably reliable network, MTBF, MUT and MDT are obtained using the formulas described in Cavers [2].

$$\text{MTBF} = \left[\sum_{i=1}^N R_i \right]^{-1} \quad (13.1)$$

$$\text{MUT} = \text{MTBF} (1 - P(F)) \quad (13.2)$$

$$\text{MDT} = \text{MTBF} \cdot P(F) \quad (13.3)$$

where $P(F)$ is defined in (8) and R_i is defined in (12). From the observation of (8) and (12), the computation for MTBF, MUT and MDT are simple once the cut-vertex sets are obtained.

As stated earlier, topology management is critical to the IMS. Using cut-vertex sets we can quantify how effectively a

network is in utilizing its resources, and the IMS can use this information to make adjustments. However, we have experienced that the computation of the cut-vertex sets is a time consuming process and further study is required in finding the faster algorithms. One other possible alternative is to find the reliable sub-optimal method to evaluate the topology of the IMS system rather than using the optimal or direct method which actually computes the unlimited size of cut-vertex sets.

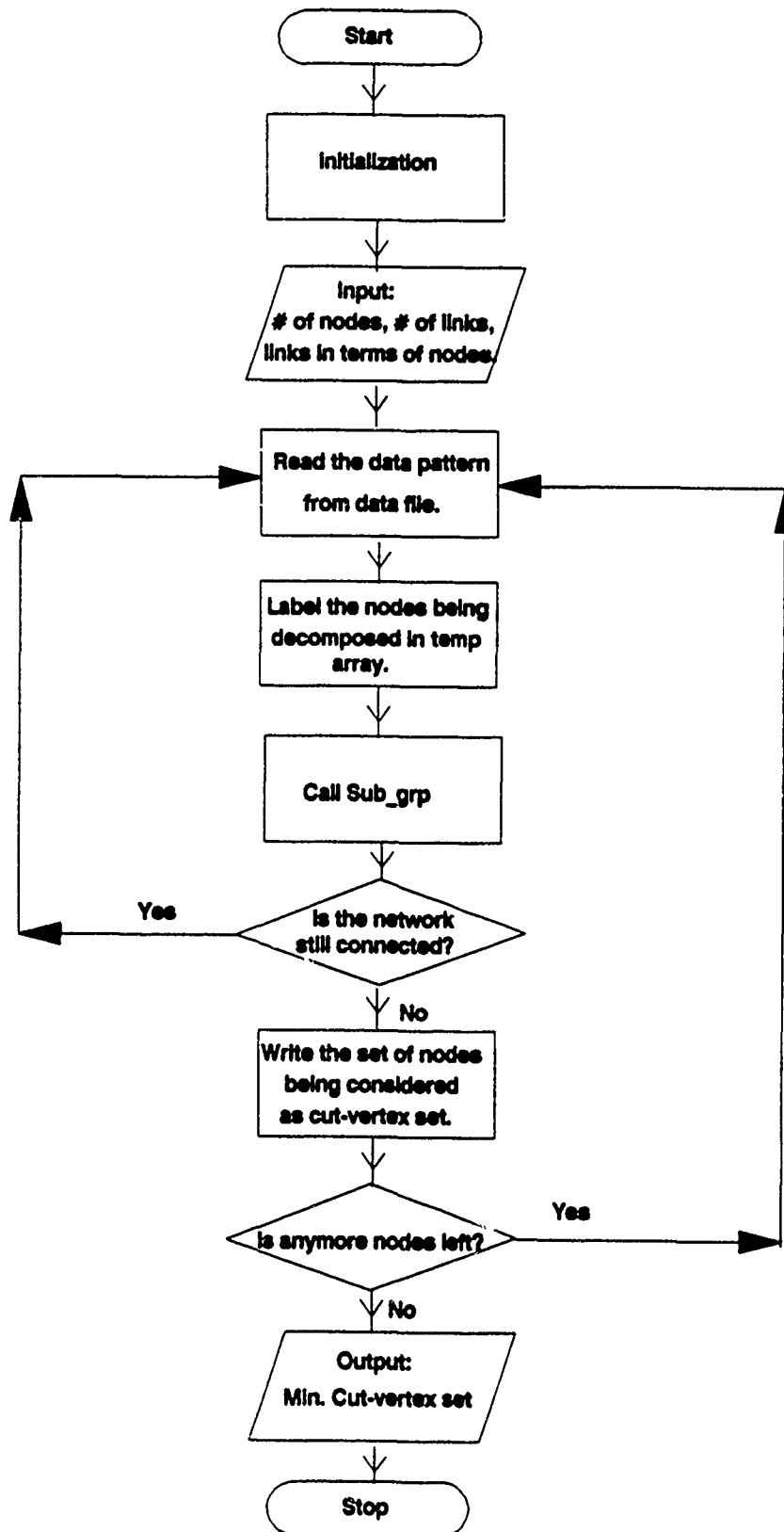
According to equation (2) and equations (9), (10) and (11), a large cut-vertex set is insignificant in terms of survivability and reliability in IMS. Therefore, we can set up a threshold (say the maximum cut-vertex size is 5 for a 20 node network or 25% of the total nodes) and compute the cut-vertex sets. If the number of nodes in the node set being considered reaches the threshold, we can just discontinue the computation process and assume the network has a reasonably high degree of survivability and reliability. This is practical implementation because we need to be particularly concerned with the cases where a network is vulnerable to a small cut-vertex set. For example, if we have a 20 node network and we know that an enemy must take out a certain 5 nodes (or less) to disjoint the network, then that knowledge is useful to the IMS to take preventive action. However, for cut-vertex sets with larger number of node elements (e.g. cut-set with 10 nodes for a 20 node network) are not useful

because it is impractical for the IMS to reconfigure such a large scale. Thus it is helpful to set a threshold value for limiting the calculation of cut-vertex sets, because the goal is to respond quickly to warnings which indicate that the survivability of an entire network is sensitive to an attack on a small set of nodes.

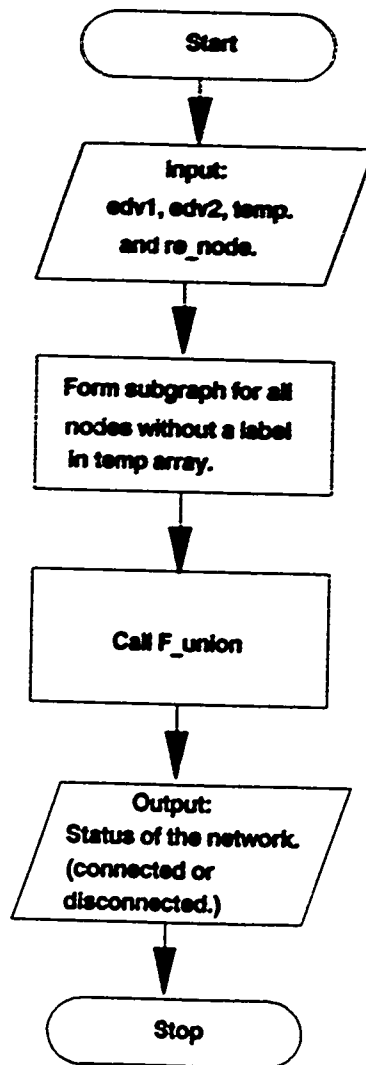
IV. Conclusion

The Communications Network Operating System II (CNOS II) program at Rome Laboratory is developing the system specifications and a prototype testbed of an Integrated Communication Network Management System (IMS). This effort is being performed by the Stanford Telecommunications, Inc. of Reston, Virginia and is scheduled for delivery in March 1992. The results of the research on cut-vertex set analysis in this paper will be integrated into the IMS prototype development effort, and will be tested within the IMS testbed to determine the effectiveness of cut-vertex set analysis in a real-time network management system. Information on cut-vertex sets can be passed to network management decision making algorithms and used to perform management/control actions to enhance network survivability. The IMS prototype is the ideal environment for evaluating the efficiency and accuracy of alternative algorithms, technologies and procedures such as cut-vertex set analysis, the results of which may lead to future applications of this work in network management.

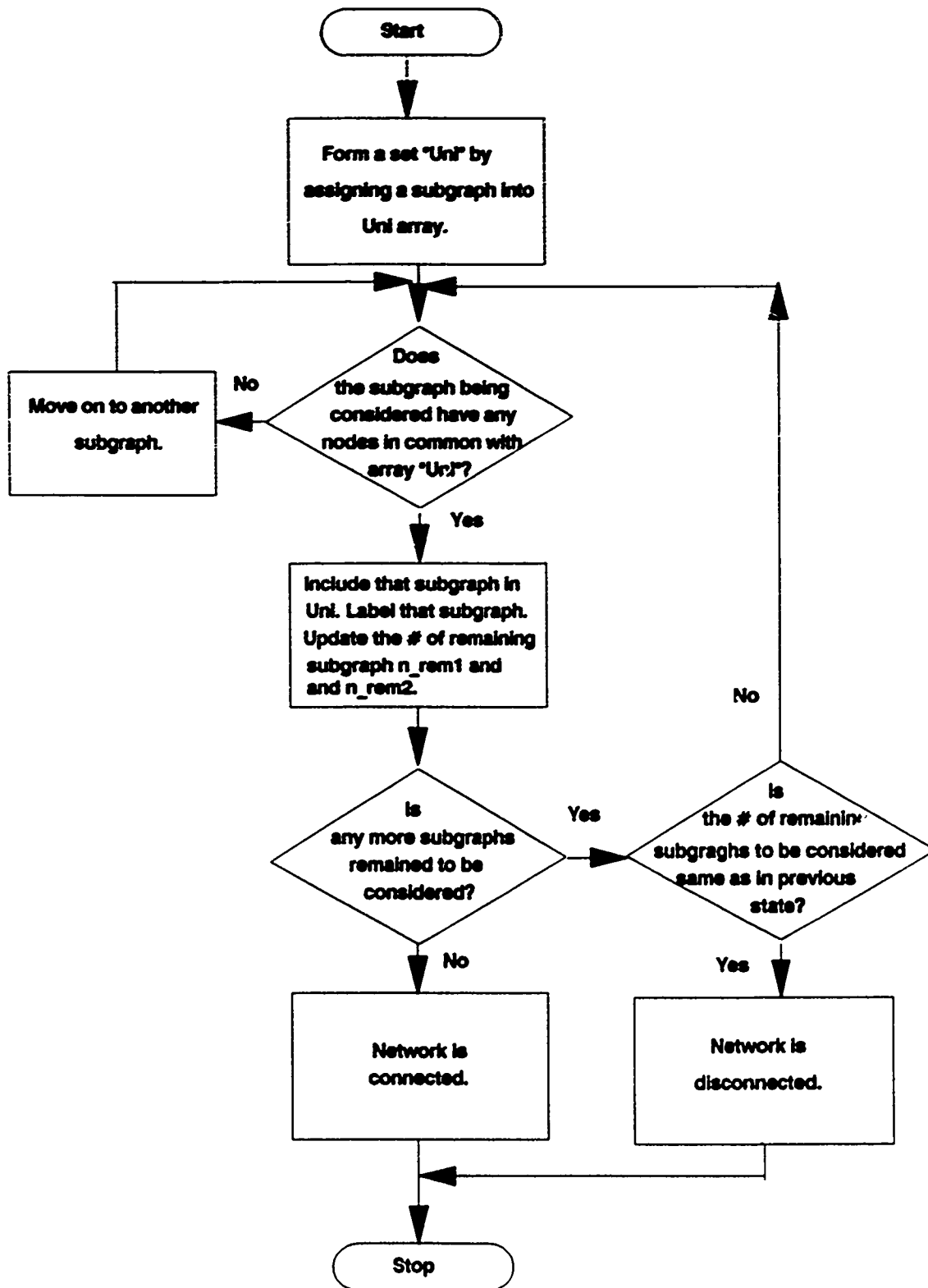
APPENDIX



Program Cut



Function Sub_grp



Function F_Unlon

REFERENCES

- [1] Soorstyn, R. R., and Frank, H., "Large-Scale Network Topological Optimization," IEEE Transactions on Communications, vol. com-25, No.1, January 1977
- [2] Cavers, J. K., "Cutset Manipulation for Communication Network Reliability Estimation," IEEE Transaction on Communications, vol. com-23, No. 6, June 1975.
- [3] Kowalchuk, N. P., " Machine Intelligence for Survivable Communications Network Management," presented in Rome Laboratory, May 1991
- [4] Newport, K. T., and Schroeder, M. A., "A Methodology for Quantifying Network Connectivity Using a Graph Theory Approach," MTR-92-78, The Mitre Corporation, Rome Laboratory, NY, August 1986.
- [5] Newport, K. T., and Schroeder, M. A., "Node and Link Topology Criteria for Survivable Network Design," MTR-9281, The Mitre Corporation, Rome Laboratory, NY, October 1986.
- [6] Newport, K. T., and Schroeder, M. A., "Netconn: A Tool for Evaluating the Topological Survivability of Communication Networks," WP-27594, The Mitre Corporation, Rome Laboratory, NY, March 1988
- [7] Newport, K. T., Whittaker, G. M., and Schroeder, M. A., "A Knowledge-based Approach to the Computation of Network Nodal Survivability," Proceeding of the 1990 IEEE Military Communications Conference, vol 3, pp 1114-1119, September

1990.

- [8] Newport, K. T., Schroeder, M. A., and Whittaker, G.M., "Techniques for Evaluating the Nodal Survivability of Large Networks," Proceeding of the 1990 IEEE Military Communications Conference, vol 3, pp 1108-1113, September 1990.
- [9] Nijenhuis, A., and Wilf, H. S., "Combinatorial Algorithms for Computers and Calculators," 2nd Edition, Academic Press, 1978
- [10] Papoulis, A., "Probability, Random Variables, and Stochastic Processes," 2nd Edition, McGraw-Hill Book Company, NY, 1984
- [11] Plotkin, M. S., and Einhom, S., "Reliability for Continuously Operating Systems," IEEE Transaction on Reliability, vol. R-14, pp.15-22, March 1965.
- [12] Stanford Telecommunications, Inc., "Service Manager of Integrated Node Final Technical Report," RADC-TR-90-198, September 1990.
- [13] Stanford Telecommunications, Inc., "Resource Manager of Integrated Node Final Technical Report," RADC-TR-90-255, October 1990.
- [14] Stanford Telecommunications, Inc., "The Integrated Communications Network Management System (CNOS II Program) Quarterly Progress Report," TB-91040, February 1991.
- [15] Stevens, J. A., and Parekh, H. P., "Communications Network Operating System I (CNOS I Program) Final

Report," Rockwell International Corp., RADC-TR-89-221,
October 1989

- [16] Syslo, M.M., Deo, N., and Kowalik, J. S., "Discrete
Optimization Algorithms with Pascal Programs," Prentice
Hall, NJ, 1983

Demonstration of a
Low Energy Electron Diffraction
System

Submitted as a Final Report
for the AFOSR/RDL Summer
Graduate Associateship Program

Yolanda J. Kime
369
September, 1991

Introduction/Statement of Problem

Over the years, one of the tasks assigned to the Reliability Physics Branch of Rome Laboratories is the investigation (and hopefully control) of electromigration in thin metal films. The atoms of a thin metal film are observed to move and migrate under the influence of electric fields: this is the process of electromigration. Electromigration can occur even in low electric fields and relatively small current densities. Electromigration can be so severe that holes, or 'voids', appear in the metal films. The number and size of the voids can increase to the point that the electric resistance of the film is many times greater than before electromigration occurred. If the thin film is part of a microelectronic chip, the chip eventually fails: clearly a problem in the domain of Reliability Physics.

The Reliability Physics Branch has several techniques available to tackle this problem. A Scanning Electron Microscope (SEM) is available and has already proven instrumental in electromigration studies. A Scanning Auger Electron Microscope contributes elemental information to the work. Soon photoemission spectroscopies will also be available to probe the electronic structure of the thin films as they grow and undergo electromigration. In order to glean structural information about the surfaces and films a Low Energy Electron Diffraction (LEED) system was proposed and

installed. It is the demonstration of this Low Energy Electron Diffraction system which is the focus of this report.

At atmospheric pressure an atomically clean surface will be covered with a layer of contaminants in less than a second. In order to maintain a clean surface long enough to do experiments (and indeed to permit the cleaning of a number of materials) LEED, like most other surface spectroscopies is performed with both the sample and the analyzer under ultra-high vacuum. The cleaning of the sample and the ultra-high vacuum (UHV) restriction are key considerations, then, in the design of appropriate sample holders and the demonstration of the LEED equipment. The development of the sample holder, the cleaning procedure for the sample to be examined for demonstration purposes, as well as the LEED technique itself will be discussed in the following sections.

The Low Energy Electron Diffraction Technique

As the name implies, LEED is a diffraction technique, much like x-ray diffraction or laser diffraction techniques used in other technologies. In any diffraction technique, the incident wave hits the sample and bounces off. Different parts of the same wave front may be reflected from different atoms (see Figure 1). The total path lengths traveled by the two parts of the wave will then be different. If the difference in the path length is equal to the peak-to-peak wavelength of the incident wave (or an integral number of wavelengths), the two scattered parts of the wave front are said to be in phase, and the two scattered parts reinforce each other (the peak of one wave adds to the peak of the second wave). This is called constructive interference. If, on the other hand, the difference in path length is equal to half of the peak-to-peak wavelength of the incident light then the peaks of one part of the scattered wave add to the troughs of the other part of the scattered wave, and the two parts cancel each other out. This is called destructive interference.

It thus becomes a question of geometry. The wave scattered in some directions will have interfered

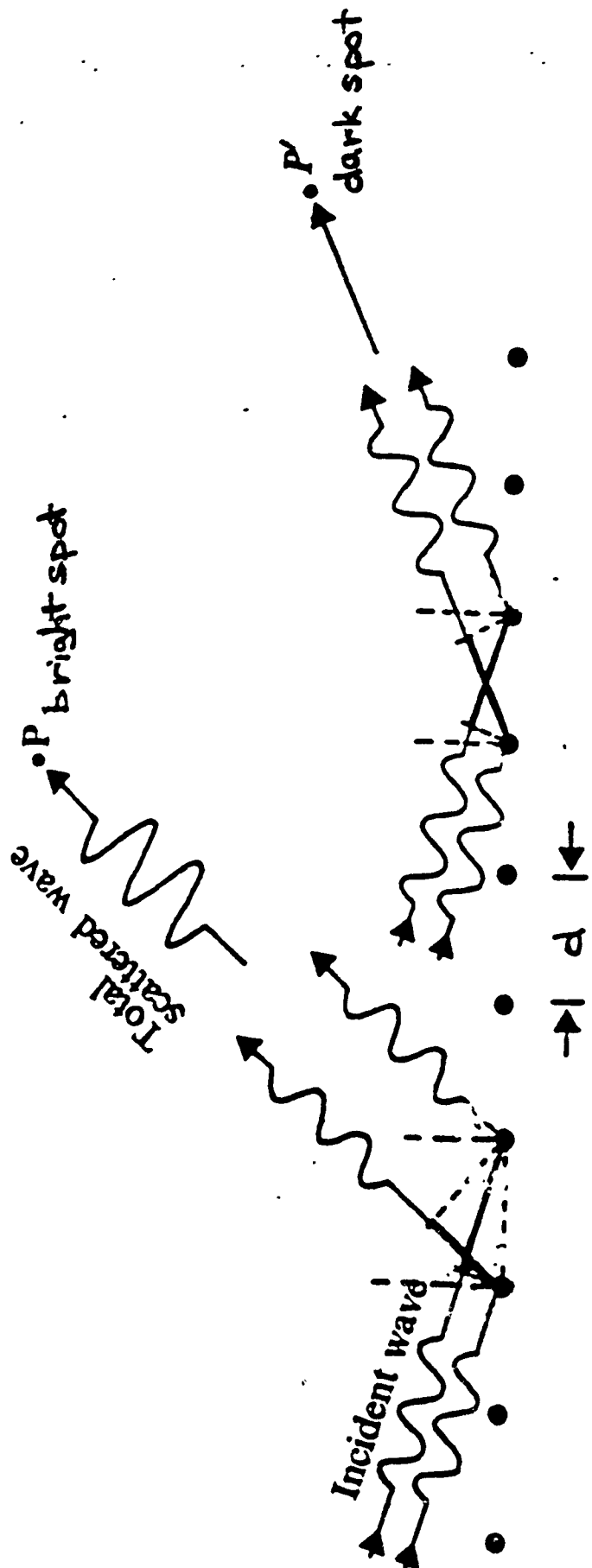


Figure 1

constructively and so the wave will be intensified, and in some directions it will have interfered destructively and the wave will be cancelled out. If the incident wave hits the surface perpendicularly, we can describe the condition for constructive interference mathematically as follows. If the distance between the atoms of the surface is d and the wave is scattered at an angle θ (measured from the surface normal) then the additional path length will be $d \sin(\theta)$. So to have constructive interference (increased intensity) we must have

$$n\lambda = d \sin(\theta)$$

where λ is the wavelength of the incident wave and n is an integer.

Although one usually doesn't think of an electron as a wave, it nonetheless behaves like a wave in a lot of experiments. Quantum mechanically the wavelength of the electron is related to its energy by

$$E = (h/\lambda)^2 / 2m$$

where h is Planck's constant and m is the mass of the electron. The energy of the electron is just the kinetic energy the electron has gained when accelerated through a voltage V . Solving this equation for the wavelength and putting in the constants, we get

$$\lambda = \sqrt{1.504/V}$$

where the voltage is in volts and the wavelength is in nanometers.

The advantage of using electrons as the incident waves instead of x-rays or visible light is that electrons do not penetrate very well into the surface under investigation. This means that electron diffraction is very surface sensitive. Electron diffraction gives information only about the top few layers of atoms, making it an ideal probe for surfaces or very thin films. When coupled with a technique that gives information about what kind of atoms are sitting on the surface, such as Auger Electron Spectroscopy, LEED can be a powerful tool in understanding the thin films used in microelectronics.

When LEED is done in a laboratory, a beam of electrons are accelerated to a chosen energy, usually less than 500 eV (electron volts) in an electron gun. The beam of electrons hits the sample (usually placed perpendicular to the beam) and it is diffracted back. The diffracted electrons are observed by using a phosphor coated screen. The phosphor emits light when hit with an electron so a bright spot appears on the screen where the electron wave interfered constructively and the screen is dark where the reflected waves cancel each other out. Usually a slight retarding

potential is put on a second screen in front of the phosphor screen so that only the perfectly diffracted electrons get through and light the screen.

In general the electron gun and the phosphor screens are purchased as a unit; the challenge is to configure the sample holder and the experimental chamber to accomodate whatever experiments the LEED is to be used for. This is the topic of the next section.

Sample Holder Design and Chamber Configuration

Although the experiment used to demonstrate the LEED system was a simple examination of a Ag(100) single crystal surface, the system is intended to be used to do electromigration studies. This imposes some restrictions on the design of the sample holder and the chamber configuration. For example, the sample must be electrically isolated from the chamber. This allows a controlled current to be passed through the sample, both to produce electromigration and to resistively heat the sample. Also, one would like to be able to separate changes in the surface due to the applied current from those changes due only to thermal processes. To this end, sample cooling and temperature measurement of the sample are also included in the design. Sample placement (and therefore sample holder design) is limited also by the electron optics of the LEED screens and the electron gun. All of these requirements must be met in a design compatible with the ultra-high vacuum required for sample cleanliness.

The sample holder is shown schematically in Figure 2. The sample holder consists of a hollow, L-shaped, stainless steel tube. Three 1 mm diameter tungsten rods run down the center of the tube and out the end through ceramic-to-metal

SAMPLE HOLDER

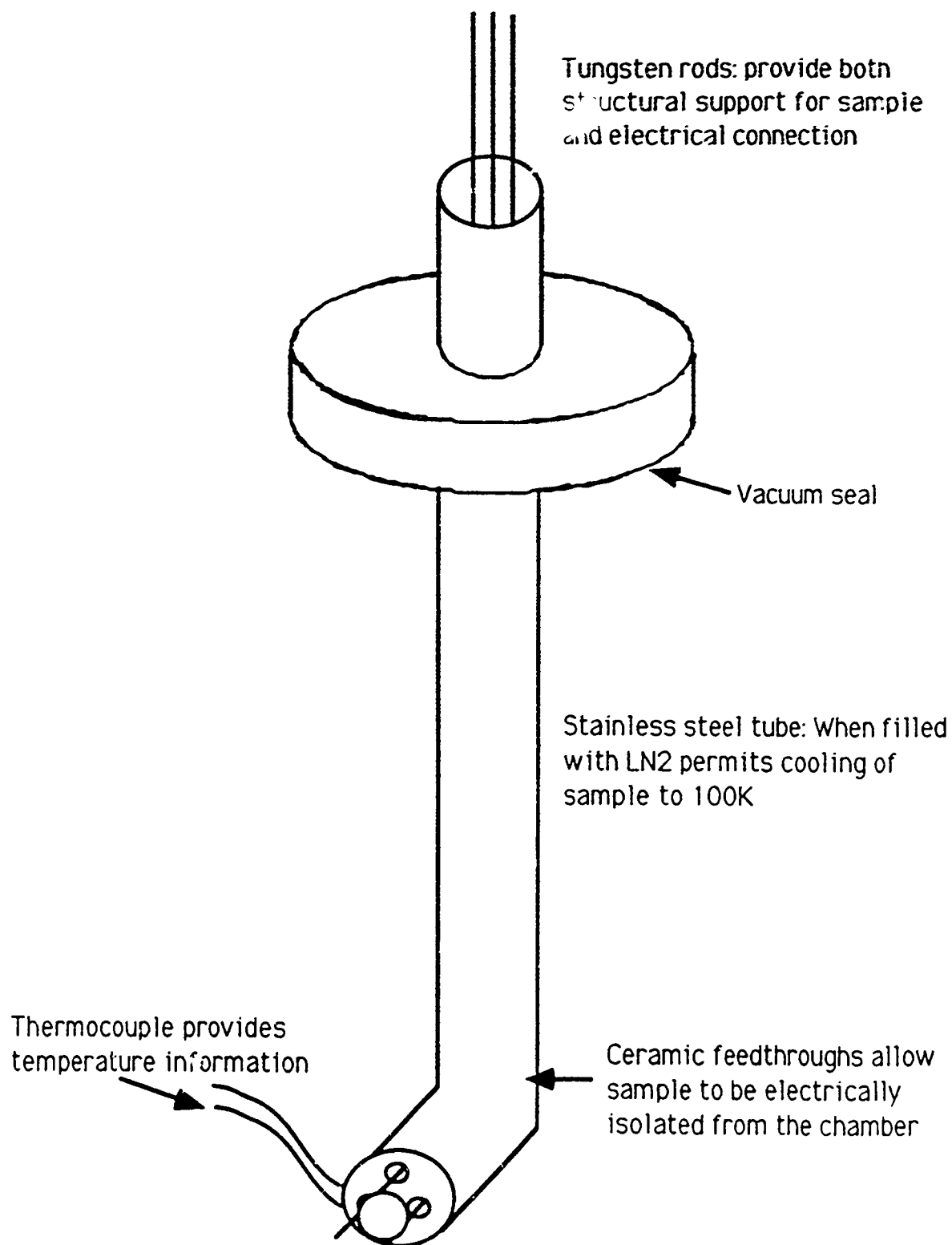


Figure 2

feedthroughs which provide electric isolation and vacuum compatibility. The sample is supported by loops of thin nickel wire which are spot welded to the tungsten rods. The sample can be cooled to 130 K (as determined by a Chromel-Alumel thermocouple) by pouring liquid nitrogen into the sample holder. For this reason, this kind of a sample holder is sometimes referred to as a "coldfinger".

The chamber configuration is also important. The chamber is pumped by an ion pump, a titanium sublimation pump, and a turbo-molecular pump which is backed by a mechanical pump. Besides the LEED screens and electron gun, the chamber is also equipped with a Varian ion gun, a mass spectrometer, and an ion gauge. The ion gauge monitors the pressure in the chamber and the mass spectrometer can be used for leak checking or Thermal Desorption Spectroscopy (TDS), another useful surface analysis technique. The ion gun is used to clean the sample.

Both the ion gun and the LEED optics are mounted in the same horizontal plane. A rotary motion feedthrough mounted on the coldfinger allows the coldfinger to be turned from facing the ion gun to facing the LEED screens.

Sample Preparation and LEED Demonstration

As a caution to the reader, I would note that appropriate sample preparation depends dramatically on the material of the sample used. The proper cleaning procedure for many materials can usually be found in the literature. The method outlined below is appropriate for the Ag single crystal, but may be quite damaging to other surfaces.

The sample used to demonstrate the LEED system was a Ag(100) single crystal, oriented by x-ray backscattering and spark cut to a thickness of about 1 mm. The crystal was polished with a series of diamond paste solutions, starting with 15 micron (u) diamond paste, followed by 8 u, 5 u, 1 u, and 0.5 u paste. The crystal was then cleaned with acetone and methanol and mounted on the sample holder. The thermocouple was mounted on the sample and the coldfinger mounted on the chamber.

The crystal was cleaned by repeated cycles of Ar⁺ sputtering with 500 eV ions and annealing to roughly 500 °C. The first sputter of each day was done with the sample at room temperature, with subsequent sputtering done at 300 °C. Typically, the sample was sputtered cold for about an hour, then sputtered warm for 45 minutes followed by a half hour anneal. This sequence of sputtering and annealing was done until a clean, well ordered surface was produced, as evidenced by the observation of LEED spots. As is usually the case, cooling the silver with liquid nitrogen after the final anneal sharpened the LEED image.

Appendix: Notes for the User

There are a number of things unique to the LEED chamber in use at Rome Labs. Following is a collection of notes which may be useful for persons operating that system.

1. For this system, the sample should be placed 0.75 inches from the front edge of the LEED screens. Great care should be taken in mounting the sample so that it is parallel to the LEED screens. Some correction of this angle is provided by the bellows surrounding the coldfinger, but gross errors cannot be corrected for.

2. If the sample is placed at the required 0.75 inches from the LEED screens, the angle between the surface normal and the outside edge of the screen is roughly 70° . If the electron accelerating voltage is tuned so that the first order spots ($n=1$) are at the edge of the screen, the separation d between the diffracting atoms on the surface may be easily obtained from

$$d = 1.064 \sqrt{(1.504/V)}$$

where V is in volts and d is in nm.

3. When removing the coldfinger from the chamber, care should be taken to break the vacuum BELOW the rotary motion feedthrough. Because of the shape of the coldfinger it cannot be slipped through the rotary motion feedthrough: if the vacuum is broken above the rotary motion feedthrough the coldfinger must be rewelded in order to replace the copper gasket above the feedthrough.

**1991 USAF-RDL SUMMER RESEARCH PROGRAM
FOR FACULTY AND GRADUATE STUDENTS**

**Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH**

FINAL REPORT

**ANALYSIS OF THE ELECTROMIGRATION INDUCED FAILURE
IN THE VLSI INTERCONNECTION COMPONENTS AND
THE MULTISECTION INTERCONNECTIONS**

Prepared by : Ashok K. Goel, Ph.D.
 Assistant Professor
 and
 Matthew M. Leipnitz
 Graduate Student

Department : Department of Electrical Engineering

University : Michigan Technological University

Research Location : Reliability Physics Laboratory
 RL/RBRP
 Rome Air Development Center
 Griffiss Air Force Base
 Rome, N.Y. 13441

USAF Researcher : Martin J. Walter

Date : July 17, 1991

Contract Number : F-49620-90-C-0076

ANALYSIS OF THE ELECTROMIGRATION-INDUCED FAILURE IN THE VLSI INTERCONNECTION COMPONENTS AND THE MULTISECTION INTERCONNECTIONS

Ashok K. Goel, Assistant Professor, and

Matthew M. Leipnitz, Graduate Student

Department of Electrical Engineering
Michigan Technological University
Houghton, MI 49931

ABSTRACT

We have carried out a first-order analysis of the electromigration induced failure effects in the various VLSI interconnection components including the multisection interconnections using the series model for failure mechanism. The Components include a straight interconnection segment, an interconnection bend, an interconnection step, an interconnection plug, an interconnection via, an interconnection overflow, a horizontal multisection interconnection, a vertical multisection interconnection, a mixed multisection interconnection, and a power/ground bus. First, by considering the effect of average flux density on the grain-boundary migration, we have reduced each interconnection component in to a series or series-parallel combination of straight segments. Then, for each of the components, we have investigated the dependence of the median-time-to-failure and the Log-normal standard deviation of the corresponding failure distribution on the various component parameters. The results can be utilized to choose optimum values of the component parameters for minimum probability of interconnection failure due to electromigration.

1. INTRODUCTION

Continuous advances in the field of VLSI are resulting in smaller integrated circuit chips having millions of interconnections that integrate the components on the chip. Interconnection failure caused by electromigration is one of the major factors responsible for lowering the effective lifetime of the chip [1-3]. Therefore, it is very important to understand the dependence of the electromigration induced interconnection failure on the various interconnection parameters. In fact, in the past, several studies have been dedicated to this effort [4-8].

In general, an interconnection line on an IC Chip consists of several components such as straight segments, bends, steps, plugs and vias. In addition, there are power and ground buses serving several logic gates on the chip. For submicron width interconnection lines, there can be sections along the line length suffering from material overflows. In this report, for the first time, we also introduce a multisection interconnection which can be designed in three possible configurations: horizontal, vertical and mixed. A multisection intersection differs from a generally employed interconnection in that a driver and its load are connected by more than one interconnection line thus providing more than one path for the current/voltage signal to flow.

In this report, we present an analysis of the electromigration-induced failure in each of the interconnection components listed above. First, using the effects of the average flux density on the grain boundary migration in the interconnections, we have derived expressions for the effective lengths, widths and thicknesses of the straight segments equivalent to each of the components. Then we have used the series model of failure mechanisms in the interconnections [9] to determine the series or series-parallel combinations of straight segments equivalent to each interconnection component. Finally, we have studied the dependence of the electromigration-induced Median-Time-to-Failure (MTF) and the standard deviation of the corresponding lognormal failure distribution (σ) on the various parameters of each interconnection component.

2. REDUCTION OF COMPONENTS INTO STRAIGHT SEGMENTS

First, we have analyzed a straight interconnection segment, shown in Figure 1(a), of length L , width W and thickness T carrying a current I at a given temperature. Then, by considering the effects of the average flux density on the grain boundary migration in each interconnection component, we have reduced it to a series-parallel combination of equivalent straight interconnection segments. The average flux density in a component was determined by using the interconnection current I and the average cross sectional area throughout the component.

The additional area in an interconnection bend of angle θ_B , shown shaded in Figure 1(b), was found equivalent to a straight segment of length L_B and width W_B given by the expressions:

$$L_B = \frac{\pi W (180 - \theta_B)}{360}$$

$$W_B = \frac{W^2 (1 + \sqrt{\tan(\theta_B/2)})}{W + L_B}$$

For a bend angle of 90 degrees, these expressions yield values in agreement with those derived by Frost and Poole [9].

An interconnection line of length L , width W and thickness T having a single step of height H and angle θ_S , shown in Figure 1(c), is equivalent to three straight segments each of width W , lengths L_{S_1} , L_{S_2} and L_{S_3} , and thicknesses T_{S_1} , T_{S_2} and T_{S_3} , respectively given by the expressions:

$$L_{S_1} = L + \frac{H}{\tan \theta_S}$$

$$T_{S_1} = T$$

$$L_{S_2} = T \cos \theta_S + \frac{H}{\sin \theta_S}$$

$$T_{S_2} = T \cos \theta_S$$

$$L_{S_3} = \frac{\pi T (180 - \theta_S) (1 - \cos \theta_S)}{720}$$

$$T_{S_3} = \frac{T^2 (1 + \cos^2 \theta_S - \sin \theta_S \cos \theta_S)}{2 L_{S_3} + T (1 - \cos \theta_S)}$$

Two straight sections of an interconnection line of total length L , width W and thickness T joined by a single plug of length H and square dimension W_P , shown in Figure 1(d), is equivalent to three straight segments of lengths L_{P_1} , L_{P_2} and L_{P_3} , widths W_{P_1} , W_{P_2} and W_{P_3} , and thicknesses T_{P_1} , T_{P_2} and T_{P_3} , respectively given by the expressions:

$$L_{P_1} = L - W_P$$

$$W_{P_1} = W$$

$$T_{P_1} = T$$

$$L_{P_2} = H$$

$$W_{P_2} = W_P$$

$$T_{P_2} = W_P$$

$$L_{P_3} = \frac{\pi}{8} (T + W_P)$$

$$W_{P_3} = W$$

$$T_{P_3} = \frac{T (W + W_P)}{L_{P_3}}$$

An interconnection line of length L , width W and thickness T having a length L_O suffering from overflow (top and end views are shown schematically in Figure 1(e)), is equivalent to two straight segments of lengths L_{O_1} and L_{O_2} , widths W_{O_1} and W_{O_2} , and thicknesses T_{O_1} , T_{O_2} given by the expressions:

$$L_{O_1} = L - L_O$$

$$T_{O_1} = T$$

$$W_{O_1} = W$$

$$L_{O_2} = L_O$$

$$T_{O_2} = \frac{-W + \sqrt{W^2 + 4 W T}}{2}$$

$$W_{O_2} = \frac{W T}{T_{O_2}}$$

Two straight sections of an interconnection line of total length L , width W and thickness T joined by a via of height H , width W_V and angle θ_V , shown in Figure 1(f), is equivalent to four straight segments each of width W , lengths L_{V_1} , L_{V_2} , L_{V_3} and L_{V_4} , and thicknesses T_{V_1} , T_{V_2} , T_{V_3} and T_{V_4} respectively given by the expressions:

$$L_{V_1} = \frac{\theta_V \pi T (1 + \cos(\theta_V))}{720}$$

$$T_{V_1} = \frac{T^2 [1 + \cos\theta_V \sin\theta_V + \cos^2 \theta_V]}{2 L_{V_1} + T [1 + \cos\theta_V]}$$

$$L_{V_2} = \frac{H}{\sin\theta_V} - T \sin\theta_V$$

$$T_{V_2} = T \cos\theta_V$$

$$L_{V_3} = \frac{T}{2 \sin \left[\tan^{-1} \left[\frac{T}{W_V} \right] \right]}$$

$$T_{V_3} = \frac{\frac{T^2}{2 \tan\theta_V} + T \left[W_V - \frac{T}{\tan\theta_V} \right]}{L_{V_3}}$$

$$L_{V_4} = L - W_V - \frac{H}{\tan\theta_V}$$

$$T_{V_i} = T$$

A horizontal multisection interconnection of length L consists of a parallel combination of N interconnections each of length L , width W and thickness T placed between two rectangular pads each of length L_{pad} and width $W_{pad} = (2N - 1)W$ on each side which are in turn connected to the interconnection driver and its load. A schematic diagram of the top view of a horizontal multisection interconnection is shown in Figure 1(g).

A vertical multisection interconnection of Length L consists of a parallel combination of N interconnections one of which is printed on top of the substrate while the others are embedded in the substrate exactly below the top section. Each section is of length L , width W and thickness T . The sections are connected to each other at the ends by conducting plugs each of length L_p . A schematic diagram of the side view of a vertical multisection interconnection is shown in Figure 1(h). A mixed multisection interconnection is formed by mixing the horizontal and vertical multisection interconnections i.e. it has a few sections printed on top of the substrate in addition to a few sections embedded in the substrate.

As shown in Figure 1(i), a power or ground bus serving N_g gates on the integrated circuit chip was modelled as a series combination of N straight segments carrying currents equal to $I, 2I, 3I, \dots, N_g I$ where I is the current in each gate.

3. CALCULATION OF MTF AND LOG-NORMAL STANDARD DEVIATION

First, for a basic conductor element of length $10\mu m$, the median-time-to-failure was found by using the expression [9]:

$$MTF = 1,523.0 \left[\frac{W T}{I \times 10^5} \right]^n \left[W - 3.07 + \frac{11.63}{W^{1.7}} \right] e^{10,740.74 E_a / T_K}$$

where I is the interconnection current in mA, n is the current density exponent, E_a is the activation energy of the interconnection material in eV, T_K is the temperature in Kelvins, W is the interconnection width in μm and T is the interconnection thickness in μm . Then, as a first approximation, the median-time-to-failure of a series combination of N elements (MTF_s) was found by using the expression:

$$\frac{1}{MTF_s} = \frac{1}{MTF_1} + \frac{1}{MTF_2} + \dots + \frac{1}{MTF_N}$$

whereas that of a parallel combination of N elements (MTF_p) was found by using the expression:

$$MTF_p = MTF_1 + MTF_2 + \dots + MTF_N$$

The lognormal standard deviation (σ) of a basic conductor element of width W (μm) was given by [9]:

$$\sigma(W) = \frac{2.192}{W^{2.625}} + 0.787$$

Then, for a straight segment of length L , it was calculated by using the expression:

$$\sigma_n = \sigma n^{-0.304}$$

where

$$n = \frac{L (\mu m)}{10}$$

4. THE PROGRAMS "EMVIC", "EMVIC-2" and "EMGRAPH"

To date, we have developed three programs called EMVIC, EMVIC-2 and EMGRAPH. Each of the programs is interactive and extremely user-friendly. EMVIC and EMVIC-2 are written in FORTRAN-77 while the graphics program called EMGRAPH is written in C.

The program EMVIC can be used to determine the MTF and Lognormal standard deviation of a straight interconnection segment (sis), interconnection bend (ib), interconnection step (is), interconnection plug (ip), interconnection via (iv), interconnection overflow (io), horizontal multisection interconnection (hmsi), vertical multisection interconnection (vmsi), mixed multisection interconnection (mmsi) and a power/ground bus (pgbus). (The symbols in parentheses were used to name the data files as explained below.) First, the user uses the default values or chooses his/her own values for the several parameters of any component listed above. For a straight segment, the parameters include its length (il), width (iw), thickness (it), temperature (temp), current (curr), current density exponent (cde) and its material's activation energy (iae). In addition to these parameters, the other components are defined by the additional parameters listed below:

INTERCONNECTION BEND: Bend Angle (ba)

INTERCONNECTION STEP: Step Height (sh), Step Angle (sa)

INTERCONNECTION PLUG: Plug Length (pl), Square Plug Dimension (pd), Plug
Material Activation Energy, Lower Level Material
Activation Energy

INTERCONNECTION VIA: Via Height (vh), Via Width (vw), Via Angle (va),
Lower Level Material Activation Energy

INTERCONNECTION OVERFLOW:

Overflow Length (ol)

HORIZONTAL MULTISECTION INTERCONNECTION:

Number of Horizontal Sections (nhs), Source/Sink Pad
Lengths (spl)

VERTICAL MULTISECTION INTERCONNECTION:

Number of Vertical Sections (nvs), Vertical Plug

Lengths (vpl), Plug Material Activation Energy

MIXED MULTISECTION INTERCONNECTION:

Number of Horizontal Sections (nhs), Number of
Embedded Vertical Sections (nvs), Vertical Plug
Lengths (vpl), Source/Sink Pad Lengths (spl), Plug
Material Activation Energy

POWER OR GROUND BUS: Number of Gates Served by the Bus (ng), Current in
Each Gate (gc)

After the user defines the component, EMVIC calculates the MTF and σ for it and displays these values on the screen. The user can choose to write the simulation results on an output file called EMVIC.OUT.

The program EMVIC-2 incorporates and extends the program EMVIC in the sense that it allows the user to study the dependence of MTF and σ for any component on its several parameters. First, the user can define the reference component. Then he/she can select a variable parameter and choose its lowest and highest values for the dependence analysis. The number of points at which the analysis is to be carried out can also be selected by the user. Then, EMVIC-2 calculates MTF and σ at evenly distributed points in the range of analysis and writes the results on an output file called EMVIC-2.OUT. It also creates data files named COMPONENT-PARAMETER.DAT one for each parameter of every component analyzed by the user. The values of COMPONENT and PARAMETER are the symbols enclosed in the parentheses above. These data files are written in a format required by the graphics program EMGRAPH. At this point, the user can use EMGRAPH to plot the results of any of the above dependence studies. EMGRAPH also allows the user to change the appearance of the plot in order to create custom plots. These plots can then be sent to a printer. The source codes of EMVIC-2 and EMGRAPH contain nearly 6,000 and 1,000 lines, respectively. The flow chart of EMVIC-2 is shown in Figure 2.

5. SIMULATION RESULTS

The program EMVIC-2 has been used to study the dependence of MTF and σ on the various parameters of each interconnection component. In the following results, current density exponent was set at 1.0. (Due to length restriction on this report, only a few results will be included here.)

First, for a straight interconnection segment, the dependence of MTF and σ on the segment width in the range 0.5-5 μm is shown in Figure 3. The relatively sharp increase in MTF and σ for widths less than nearly 2 μm is due to the so-called "bamboo" effect [4,5]. For an interconnection bend, the dependence of MTF and σ on the bend angle in the range 10-150 degrees is shown in Figure 4. This figure shows that a bend angle of nearly 50 degrees results in the lowest value of MTF. For an interconnection step, the dependence of MTF and σ on the step angle in the range 90.1-160 degrees is shown in Figure 5. This figure shows that MTF decreases rapidly as the step angle approaches 90 degrees. This is because of the gradual thinning of the material at the step. For a horizontal multisection interconnection, the dependence of MTF and σ on the number of horizontal sections in the range 1-5 is shown in Figure 6. This figure shows that MTF varies nearly as n^2 where n is the number of sections. This is because the current density in each section is nearly $(1/n)$ of that in the original single-section interconnection and further because all sections must fail before the interconnection fails completely. Similar dependence on the number of vertical sections was observed for the vertical multisection interconnection as shown in Figure 7. However, compared to the horizontal configuration, vertical multisection interconnection offers the advantage that it does not require any additional space on the chip. For a 1,000 μm long power or ground bus serving 100 identical gates, the dependence of MTF and σ on the current in each gate in the range 0.1-1 mA is shown in Figure 8. This figure shows that increasing the gate currents results in lower values of MTF for the bus, as expected.

6. SUMMARY AND CONCLUSIONS

To summarize, we have carried out a first-order analysis of the electromigration induced failure effects in the several interconnection components. First, each component has been reduced into a series-parallel combination of equivalent straight interconnection segments and, then, the series model of failure mechanism has been used to determine the MTF and the corresponding lognormal standard deviation (σ) for each component. The algorithms have been used to study the dependence of MTF and σ on the various parameters of each interconnection component. Though, due to the approximations inherent in the series model, the analysis presented in this report gives approximate results, yet these can be used to draw important conclusions regarding the optimization of the various interconnection components.

ACKNOWLEDGEMENTS

First, we like to thank the United States Air Force Systems Command, the Air Force Office of Scientific Research and the Reliability Physics Laboratory of the Rome Air Development Center at the Griffis Air Force Base for sponsoring this research. We also like to thank several individuals who helped in making this experience truly rewarding and enriching for us. First, we are grateful to our technical focal point Mr. Martin Walter for his constant encouragement, several useful discussions and for providing a very constructive and enjoyable environment for carrying out this research. We are also grateful to Robert Hillman, Alfred Tamburrino and Mark Pronobis for taking interest in our work and for their helpful suggestions. We also like to thank Gary Moore and other staff members of the Research and Development Laboratories for their help with the several managerial aspects of this program.

BIBLIOGRAPHY

- [1] J.R. Black, "Physics of Electromigration," Proc. 12th Annual Reliab. Phys. Symp., pp. 142-149, 1974.
- [2] P.B. Gbate, "Electromigration-Induced Failures in VLSI Interconnects," Proc. 20th Int. Reliab. Phys. Symp., pp. 292-299, 1982.
- [3] D.J. LaCombe and E.L. Parks, "The Distribution of Electromigration Failures," Proc. 24th Int. Reliab. Phys. Symp., pp. 1-6, 1986.
- [4] E. Kinsborn, "A Model for the Width Dependence of Electromigration Lifetimes in Aluminum Thin-Film Stripes," Appl. Phys. Lett., Vol. 36, pp. 968-970, 1980.
- [5] S. Vaidya, T.T. Sheng and A.K. Sinha, "Linewidth Dependence of Electromigration in Evaporated Al-0.5% Cu," Appl. Phys. Lett., Vol. 36, pp. 464-466, 1980.
- [6] J. Cho and C.V. Thompson, "Grain Size Dependence of Electromigration-Induced Failures in Narrow Interconnects," Appl. Phys. Lett., Vol. 54, No. 25, pp. 2577-2579, 1989.
- [7] Y.E. Strausser, B.L. Euzent, R.C. Smith, B.M. Tracy and K.Wu, "The Effect of Metal Film Topography and Lithography on Grain Size Distributions and on Electromigration Performance," Proc. Int. Reliab. Phys. Symp., pp. 140-144, 1987.
- [8] A.S. Oates, "Step Spacing Effects on Electromigration," Proc. Int. Reliab. Phys. Symp., pp. 20-24, 1990.
- [9] D.F. Frost and K.F. Poole, "A Method for Predicting VLSI-Device Reliability Using Series Models for Failure Mechanisms," IEEE Trans. Reliab., Vol. R-36, No. 2, pp. 234-242, 1987.

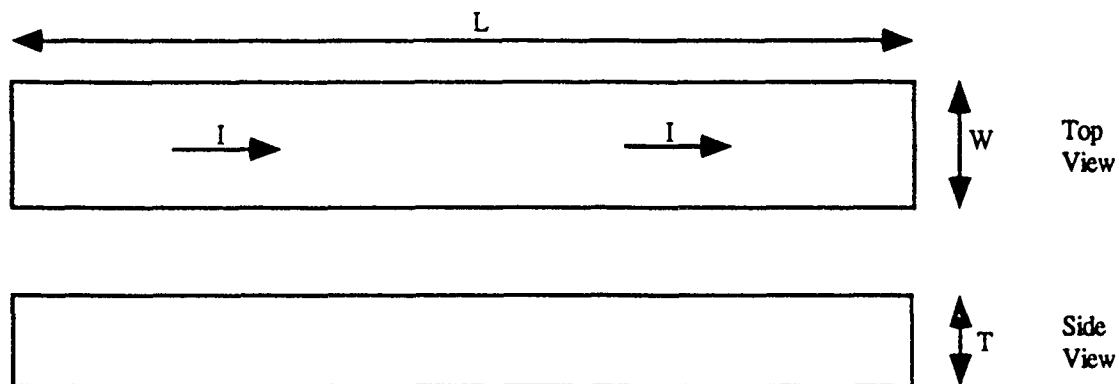


Figure 1(a): Schematic diagram of a Straight Interconnection Segment

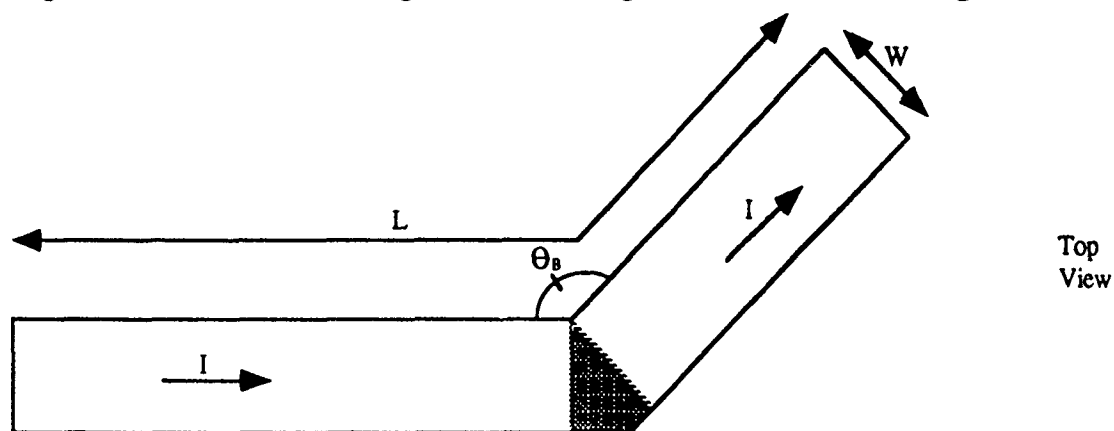


Figure 1(b): Schematic diagram of an Interconnection Bend

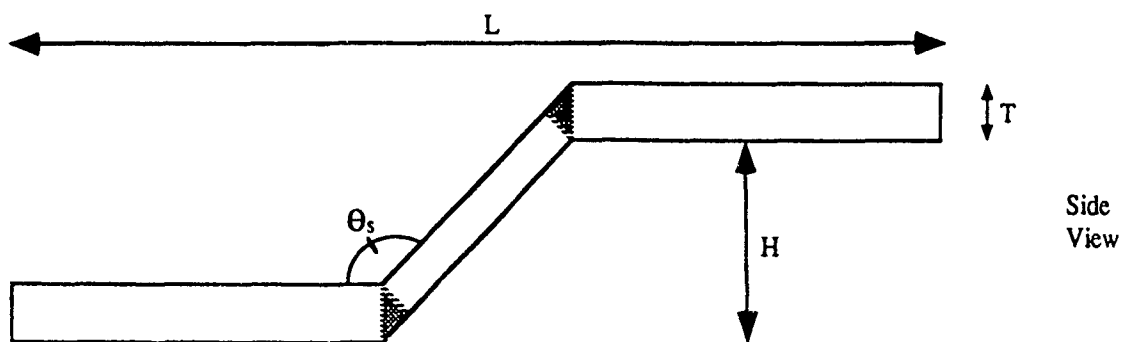


Figure 1(c) Schematic diagram of an Interconnection Step

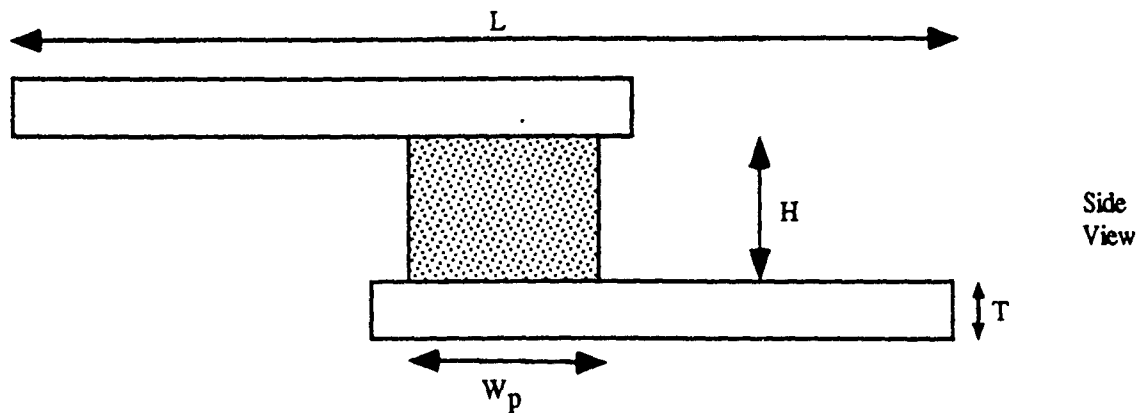


Figure 1(d): Schematic diagram of an Interconnection Plug

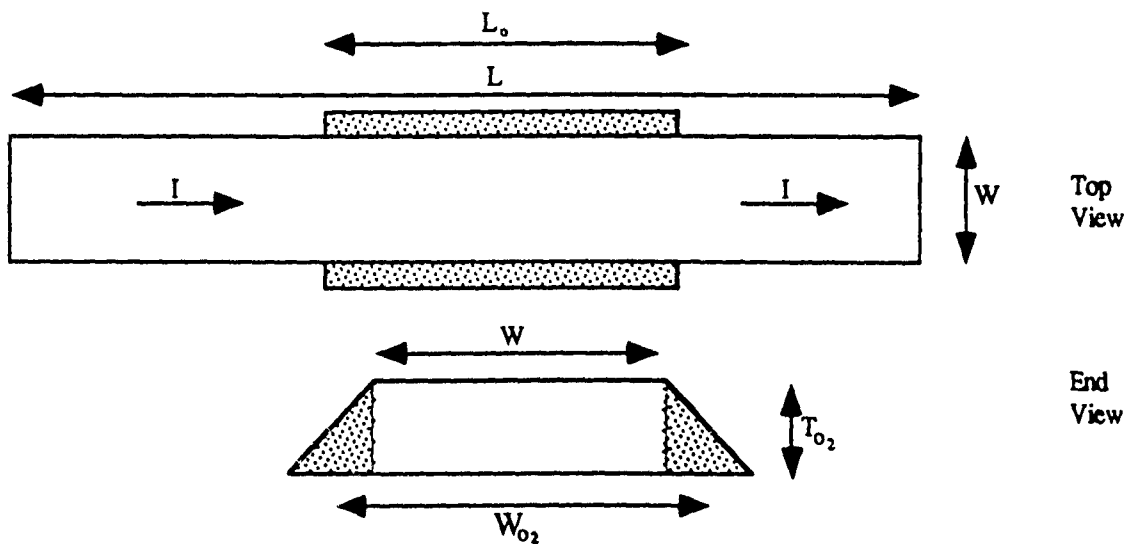


Figure 1(e): Schematic diagram of an Interconnection Overflow

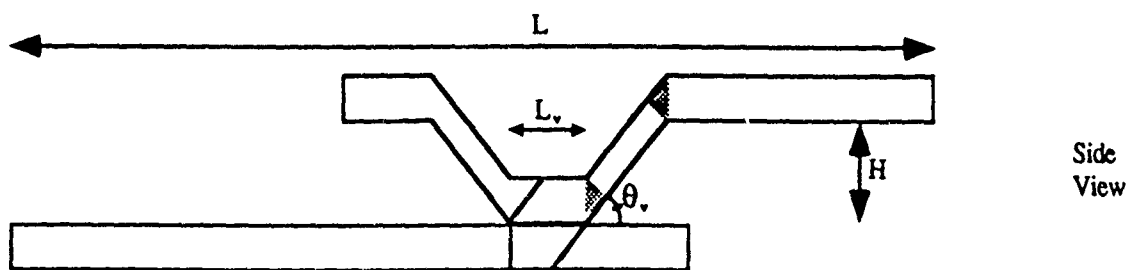


Figure 1(f): Schematic diagram of an Interconnection Via

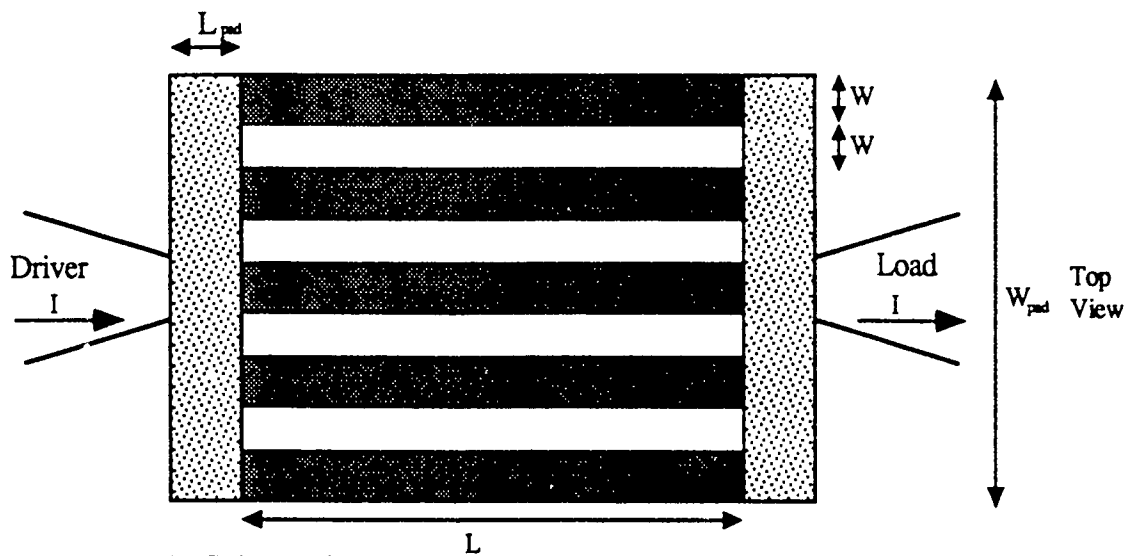


Figure 1(g): Schematic diagram of a Horizontal Multisection Interconnection

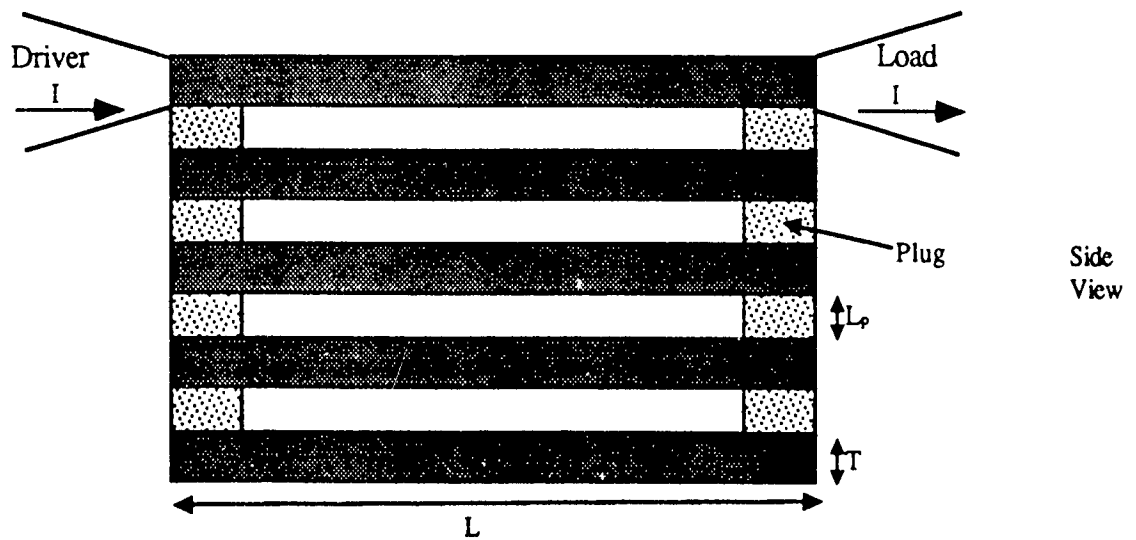


Figure 1(h): Schematic diagram of a Vertical Multisection Interconnection

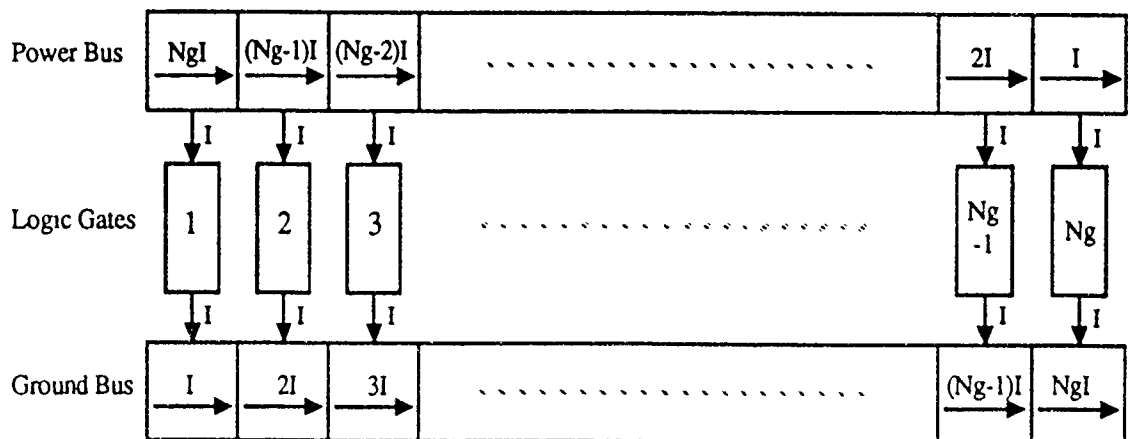
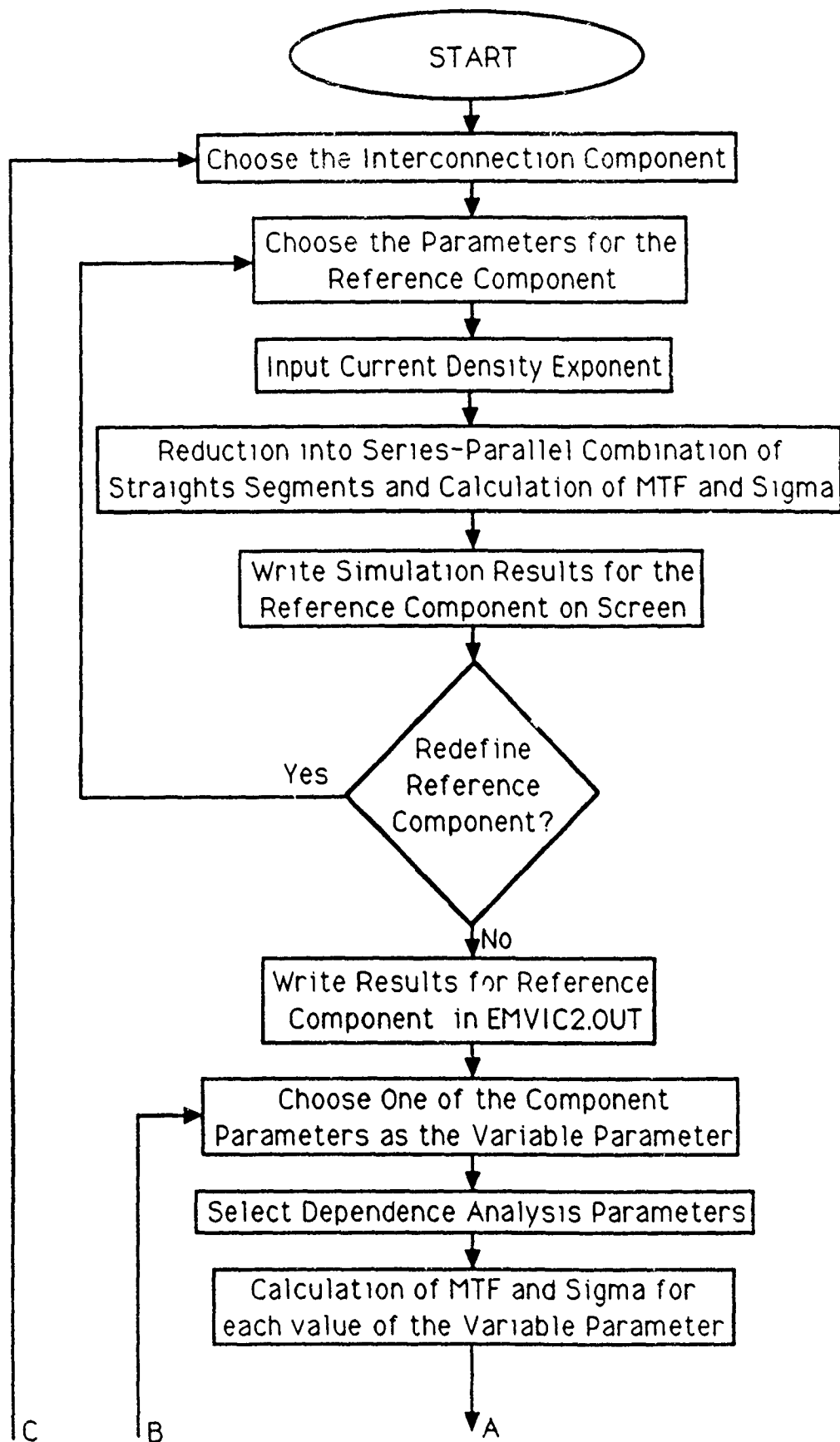


Figure 1(i): Schematic diagram of the Power and Ground Buses Serving N_g Gates on the Chip



Choose the Interconnection Component

Choose the Parameters for the Reference Component

Input Current Density Exponent

Reduction into Series-Parallel Combination of Straights Segments and Calculation of MTF and Sigma

Write Simulation Results for the Reference Component on Screen

```

graph TD
    A[ ] -- Yes --> B[Redefine Reference Component?]
    B --> C[ ]
    style A fill:none,stroke:none
    style C fill:none,stroke:none

```

↓ No

Write Results for Reference
Component in EMVIC2.OUT

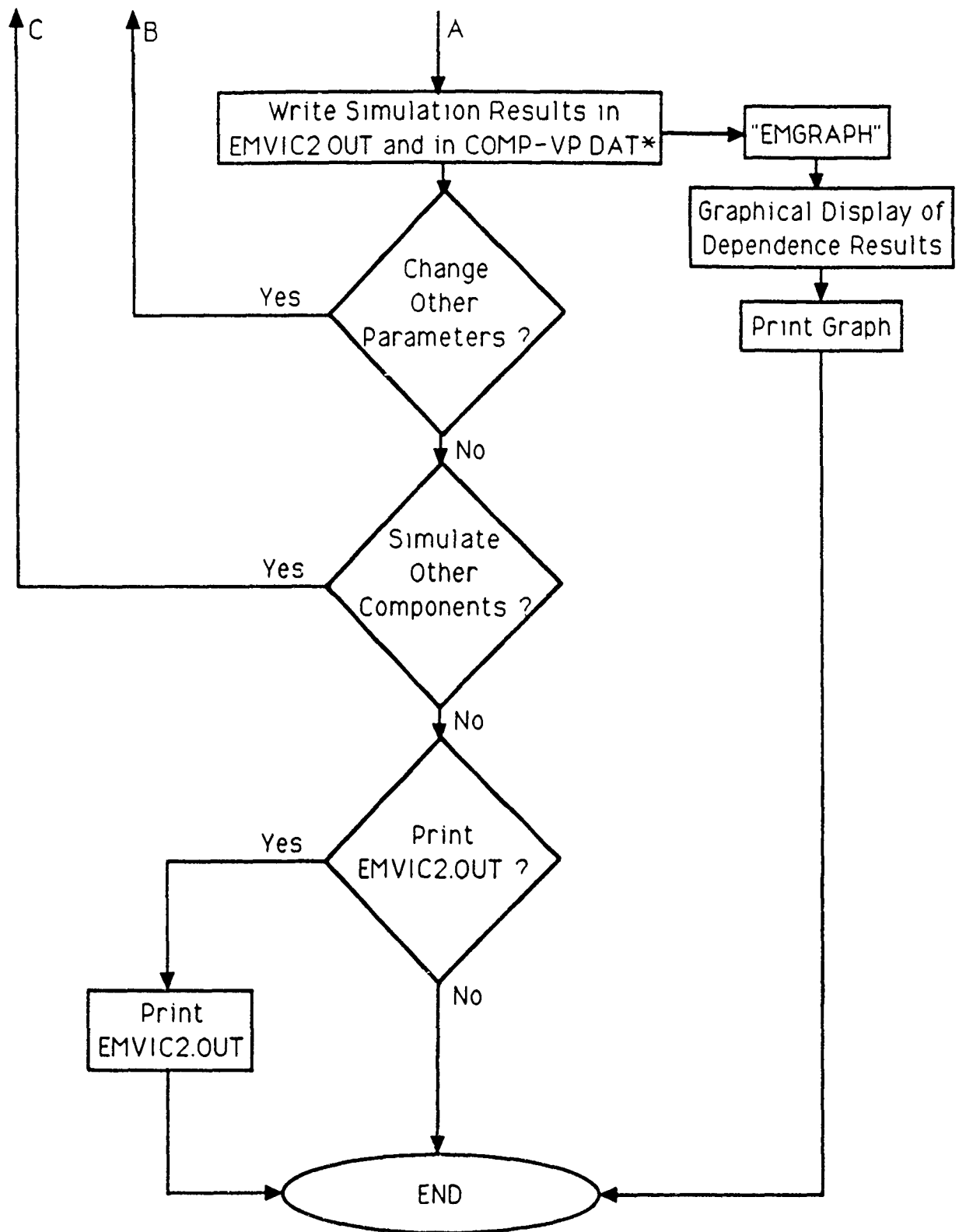
Choose One of the Component Parameters as the Variable Parameter

Select Dependence Analysis Parameters

Calculation of MTF and Sigma for each value of the Variable Parameter

A

(Continued on Next Page)



(*NOTE: The file COMP-VP.DAT stands for Component Name-Variable Parameter Name.DAT)

Fig. 2: Flow Chart of EMVIC-2

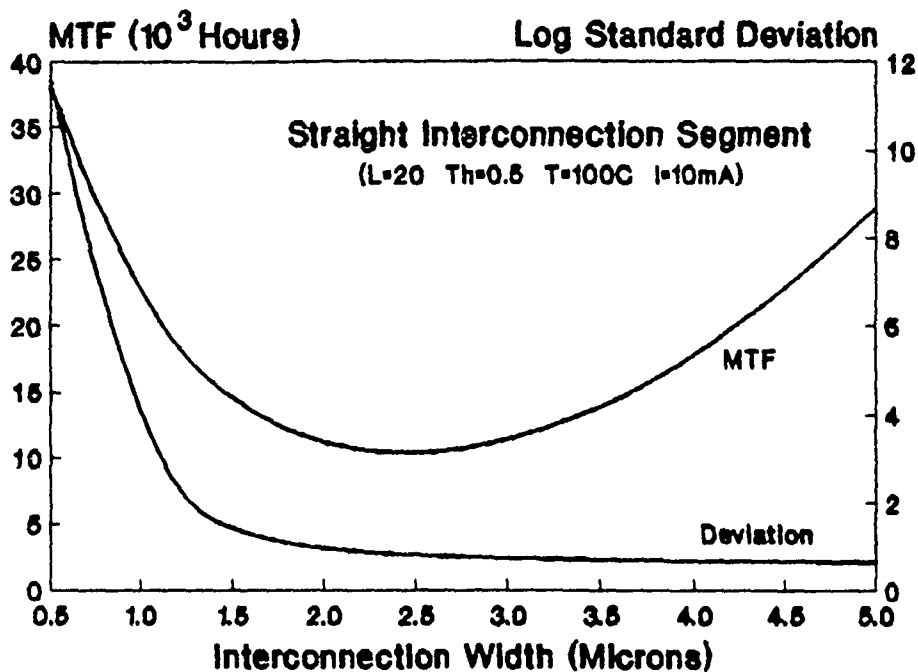


Figure 3: MTF and Deviation vs. Width for SIS

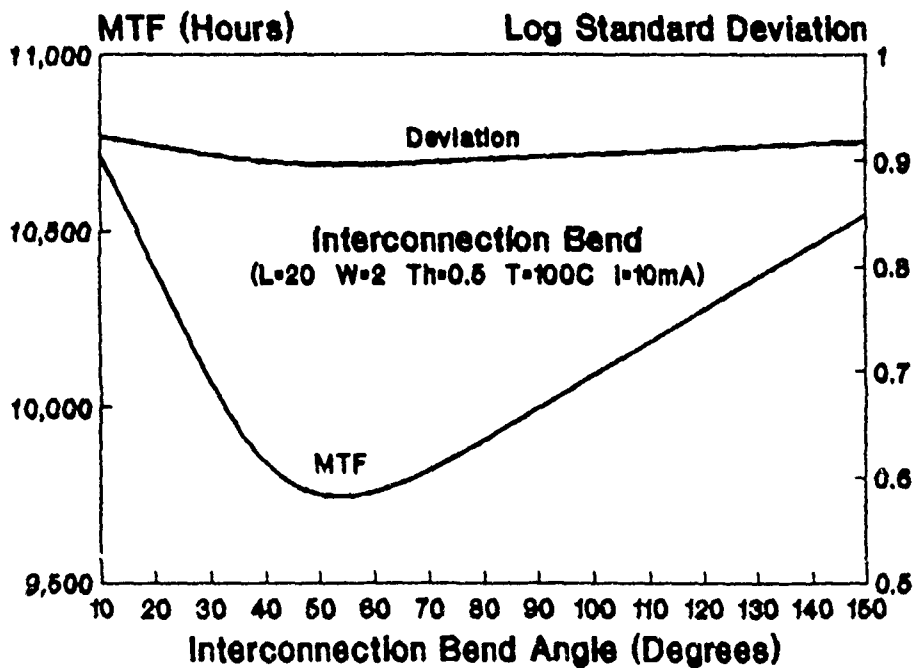


Figure 4: MTF and Deviation vs. Angle for a Bend

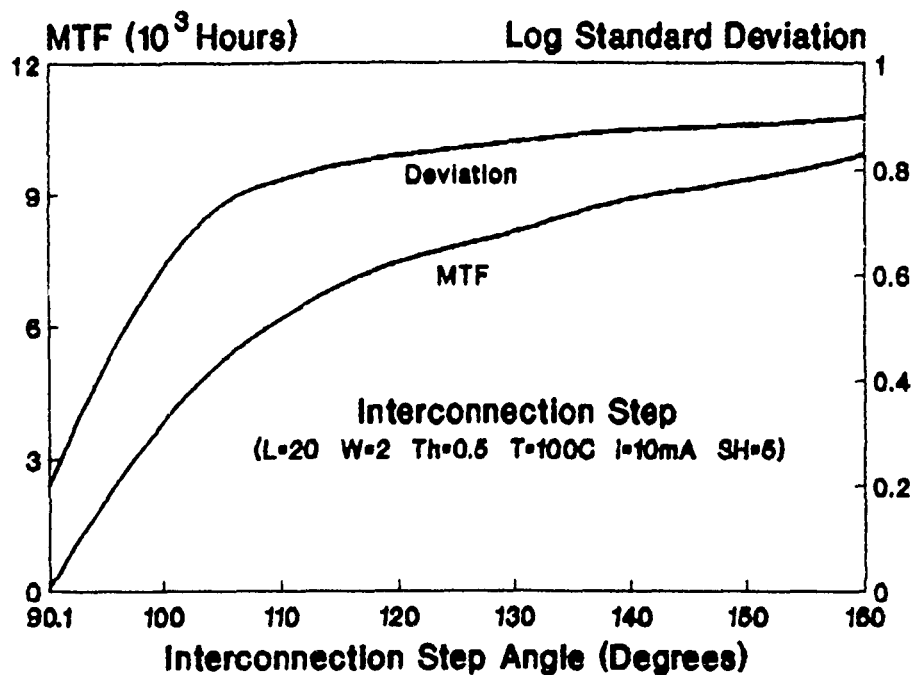


Figure 5: MTF and Deviation vs. Angle for a Step

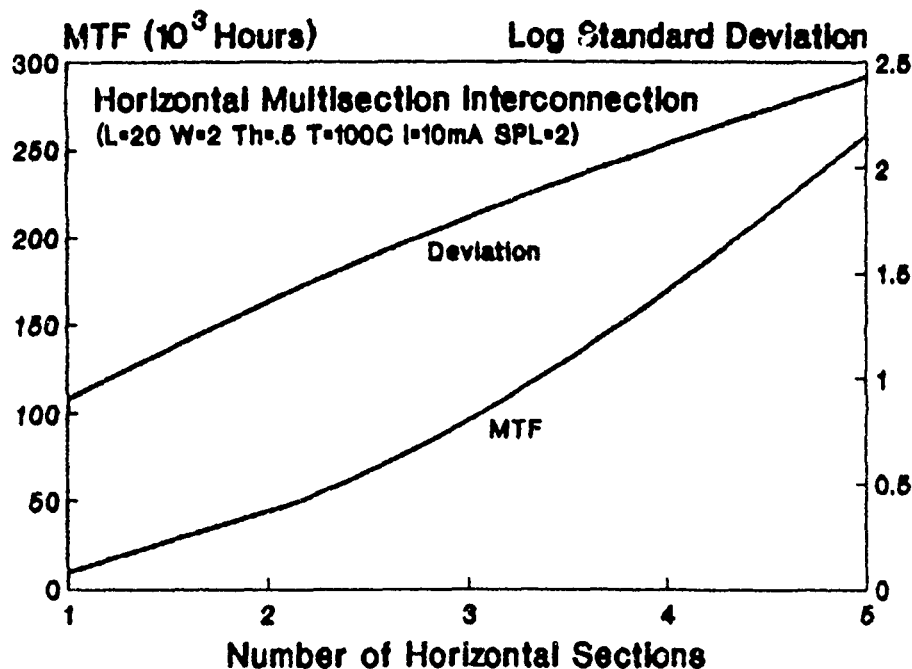


Figure 6: MTF and Sigma vs. No. of Sections for HMSI

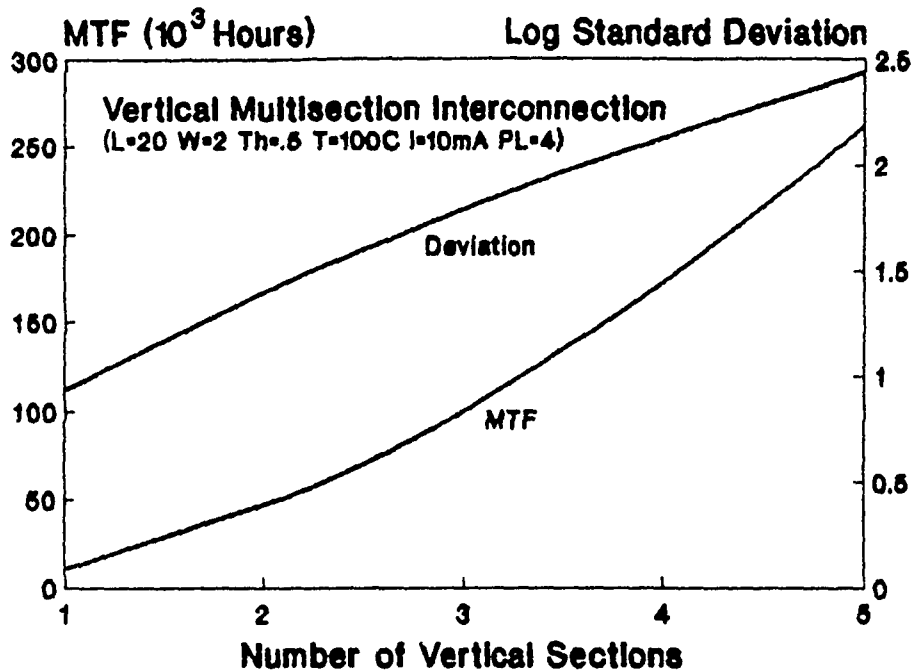


Figure 7: MTF and Sigma vs. No. of Sections for VMSI

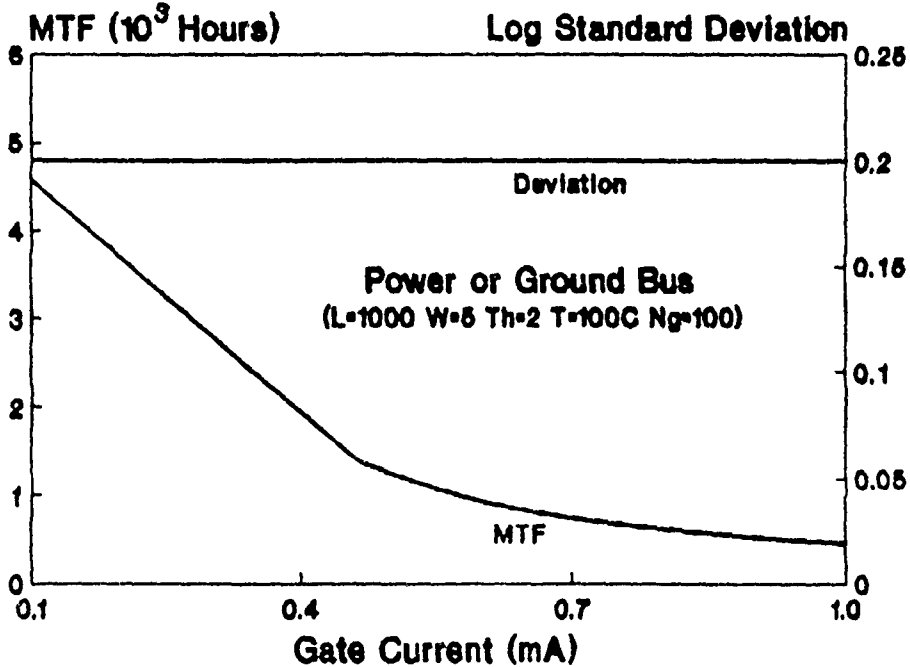


Figure 8: MTF and Sigma vs. Gate Current for a P/G Bus

WINDOWING COMPARISON PROJECT:
THE EFFECT OF WINDOW SHAPE AND SIZE ON
PHONEME IDENTIFIABILITY

Beth L. Losiewicz, MA

ABSTRACT

A variety of window sizes (4, 8 and 16 ms) and shapes (Exponential, Gaussian, Hanning, Triangular, and Rectangular) were investigated for their effect on phoneme identifiability. FFT spectrograms were created with every window size/shape combination for fricatives, stops and vowels, to see what effect different windowing parameters had on the identifying characteristics of different phonemes. For vowels, the 8 ms Hanning did the best job. For fricatives the 8 and 16 ms Gaussian and Triangular windows were best. For stops, the 4 and 8 ms Hanning window did the best job.

INTRODUCTION

For several decades, speech researchers and engineers have, with limited success, attempted to develop a machine that can recognize human speech. Assuming that humans comprehend speech through a process of segmenting speech into its constituent phonemes, researchers have attempted to isolate the characteristic signature of each phoneme using spectrograms to represent the time, frequency and amplitude characteristics of the speech signal. Today, spectrograms are usually generated by computer from a digitized speech signal, using a Fast Fourier Transform (FFT).

A variety of window functions have been developed to be used with FFT spectrograms, in an attempt to reduce the artifacts that

result from application of an FFT to a finite signal (Harris, 1978; Rhody, et. al., 1985). The problem of speech recognition has nonetheless proven relatively intractable. In the most successful machine recognition algorithms, speech signal cues have been supplemented with lexical or syntactic information, or speaker variables have been carefully controlled. Researchers continue to search for better ways to represent the speech signal (e.g. Lineback, 1984; Nathan and Silverman, 1990; Rhody, Houde, Parkins and Dianat, 1985) with the goal of improving machine speech recognition.

In this study a variety of window shapes and sizes are analyzed to see if they interact with phoneme type. That is, does any particular window shape and/or size give better information about one class of phoneme, while a different ones gives better information about a different class? Conventional speech recognition paradigms have confined themselves to a single window size/shape for the analysis of all phoneme types. But what if different phoneme types were better represented with different window sizes, or shapes? If so, multiple FFTs could be run in parallel on the speech signal, and the window best suited to each phoneme could be used in the analysis of that phoneme.

Since time information is critical in the identification of stop consonants, for example, we might expect short windows to produce the best spectrograms for their analysis, as short windows have better time resolution. On the other hand, longer windows give better frequency resolution, so we would expect them to give better information for vowel identification. In addition, longer windows

for vowels should make the resulting spectrogram less sensitive to window placement, as averaging will occur over several pitch pulses (Graff and Lundgren, 1987). With the increased computational power of modern computing devices, it is feasible to run parallel FFTs on multiple window shapes and lengths, which could then be used differentially in the analysis of different classes of phonemes.

An important caveat to the investigation reported here must be noted. The analysis of the spectrograms in this study was done visually. Presumably the human eye picks up, from a spectrogram, essentially the same information that a speech recognition machine does from the speech signal. But this is an assumption, and undoubtedly not entirely accurate (cf. Kent, 1983). Further, the analysis was carried out by a single researcher, which increases the possibility of subjective bias. As a result, this study is intended merely as a pilot, to target like window size/shape combinations, which could then be tested in machine recognition paradigms.

METHOD

Speech Data

Speech (sentence) files were selected from the TIMIT data base to obtain a representative sampling of phonemes from different phoneme classes. The fricatives chosen were /F/, /V/, /S/, and /Z/. The vowels chosen were /AA/, /AE/, /IY/, /OW/, and /UX/. All six stop consonants were analyzed. The TIMIT sentence files used, and the phonemes selected from them, are listed in Appendix A.

Phoneme samples were isolated for analysis by cutting the longer TIMIT digitized speech signal to standardized lengths. The

length of each phoneme file within a phoneme class was identical (except for the few cases noted below in which a longer speech signal was needed to capture the phoneme and its context). For fricatives, each phoneme file was 200 ms long. For vowels, each file was 200 ms long (except for /OW/ in DON'T, for the female speaker (100 ms); and /AE/ in RAG, which was 250 ms). For stops, each speech signal was 100 ms long (except /K/ in CARRY, which was 200 ms).

One phoneme file for a male, and one for a female speaker was created for each of two examples of each phoneme. Each of the two examples was taken from a different phonetic context. Thus, for example, four phoneme files were made for the phoneme /D/. A male-, and a female-speaker phoneme file for /D/ were created from the phrase "YOUR DARK;" and a male and female phoneme file were created from the phrase "DON'T ASK."

All of the TIMIT data files were originally digitized with a 16 kHz sampling rate, and 16 bit quantization.

Equipment

Spectrograms were generated using ESPRIT software (Explorer Speech Processing from Rochester Institute of Technology) adapted for a Texas Instruments MicroExplorer. The size of the sample windows always equalled the FFT point size, and a 0.94 preemphasis (6 db per octave lift) was applied to each window prior to the FFT. A floating point FFT was used with Hanning and Rectangular windows, and an integer FFT was used with the remaining window types. Spectrograms were produced in log scale, in a monotone gray-scale

dither pattern, and printed on an Apple LaserWriter IINTX with 72 DPI print resolution.

Procedures

An FFT spectrogram was produced for each phoneme file, using 4 ms (64 samples), 8 ms (128 samples), and 16 ms (256 samples) windows factorially combined with Exponential (E), Gaussian (G), Triangular (T), Hanning (H) and Rectangular (R) windows (For a description of the G, H, T and R windows, see Harris, 1978. For the E window, see Rhody, et. al., 1985. Windows were overlapped 50%.

For analysis, spectrograms were grouped by phoneme, sex of speaker, phoneme context (i.e. the phrase from which the phoneme was taken), window size, and window shape. Spectrograms were then visually compared in the following manner. Within a given window size, the best window shape was chosen for each phoneme exemplar. (Recall there were four exemplars for each phoneme - male/female X 2 phonetic contexts.) "Best" was defined as the spectrogram that exhibited the clearest critical feature information for that phoneme class. For vowels, clarity and separation of formants were the critical comparison features. For fricatives, the overall strength of the frequency components, frequency resolution at the transition from the vowel, and the presence (or absence) of low frequency energy for voiced (or unvoiced) fricatives were defined as the critical features. For stops, overall strength of the signal, clarity of transition to the following vowel, and time resolution were the critical features by which "best" was determined.

After the "best" window shape was chosen for each of the exemplars of one phoneme, results were compiled across exemplars to yield a "best" window shape for that phoneme/window size. If the "best" window shape was different for male than for female speakers, this was noted. The same phoneme files were then analyzed in the same manner for the other two window sizes. Finally, the "best" spectrograms (chosen within window size) were compared across window sizes to determine the "best" overall window for that phoneme. This procedure was repeated for each phoneme, and for each class of phonemes, allowing both specific ((to sex of speaker, phoneme context, etc.) and general (phoneme, and phoneme class-level) information to be compiled.

Results and Discussion

For brevity, window shapes will be referred to by initial letter only (e.g. Exponential window = E, etc.). Phoneme exemplars will be referred to by sex of speaker, and word context (e.g. M RAG = "rag" spoken by a male speaker.) Since all spectrograms for a given phoneme exemplar were created from a single speech signal file, all within-exemplar variation is attributable to the manipulated variation in window shape and size.

Since a variety of overlapping parameters were considered in the analysis, comment will be made on the effect of each of them: window shape, window size, sex of speaker and phoneme category. (A list of "best" windows for each phoneme exemplar is listed in Appendix B.)

Window Shape

The Triangular, Gaussian and Hanning windows gave remarkably similar results, except that the Hanning window appeared to slightly strengthen regions of strong energy and attenuate regions of lower energy in comparison with the other two. For vowels, this resulted in a clearer formant separation in the Hanning than in the other two windows. For fricatives, however, this attenuation of regions of lower energy was a drawback, since it often completely deleted a weak F0, or weak high frequency energy in a weak fricative.

The Triangular and Gaussian windows seldom differed in any essentials. Often it took careful scrutiny to verify that the spectrograms they produced were indeed different. They were equal to, or slightly inferior to, the Hanning window, except in the case of weak fricatives. For weak fricatives, the Triangular and Gaussian windows performed better than the Hanning window because of the latter's tendency to attenuate weak energy components.

Contrary to the claim of Rhody, et. al (1985), the Exponential window * performed very poorly in almost all cases, being only slightly better than the Rectangular window. Both of these windows were consistently "noisier" than the others, but might be useful if it is suspected that very weak, but important, frequency information is present in the speech signal but not detected by the other window

* But note that their results are reported for a modified Exponential window, whereas in this study a non-modified Exponential window was used. See Appendix C for a plot of the Exponential window used in this study.

types. Thus they might be useful as a double-check for voicing (F0) and high energy frequency information in very weak fricatives.

Window Size:

As expected, there was a trade off between time and frequency information as a result of window size. Overall, the 4 ms window did best only for very short stop bursts. For other phoneme categories, the 4 ms window rarely outperformed the larger windows for more than a few scattered exemplars. The 8 ms window was probably the best overall, giving much better results for fricatives, and slightly better results for vowels than the 16 ms window, in the exemplars analyzed.

In general it did not appear that window size interacted with window shape in any significant way. That is, on the whole, choice of best window shape did not change across window sizes, nor vicaversa.

The analysis of window size did yield two surprises. First of all, it appeared that, in the case of fricatives, window size differentially affected the strength of the representation of high versus low frequency energy. While the 4 ms window usually gave the most information about high frequency energy, the 16 ms window was more apt to pick up F0 information.

The second unexpected result was that the 8 ms window did a slightly better job for vowels than the 16 ms window. Possibly this is because 8 ms is closer to the average pitch period length.

After the planned analysis was completed, spectrograms were created with a 64 ms Hanning window for the phoneme exemplars

which gave the best (M APTITUDE) and worst (M DARK) vowel spectrograms in this study. This longer window was expected to be less sensitive to placement variation since it spans several pitch periods. The 64 ms window gave clear formant structure for both exemplars. In the best exemplar, M APTITUDE, the 64 ms window gave much weaker formant information than the 16 ms window did. However, for M DARK, which was unreadable in the 16 ms window, the 64 ms window gave crisp, clear formant structure.

Thus it appears that the smaller windows give adequate results when formants are well separated, but a larger window is needed to resolve formants that are quite close in frequency. However, it is also possible that the poor frequency resolution for M DARK in this study was an artefact of window placement. Further investigation of the effect of much larger windows is necessary before firm conclusions can be drawn.

Sex of Speaker:

Somewhat surprisingly, sex of speaker had little consistent effect on which window size or shape yielded the "best" window. This was especially true for fricatives and stops. For vowels, although the 8 ms window had an edge over the other two window sizes for both male and female speakers, the "second best" window was the 16 ms for males, but the 4 ms window for females. Presumably this would be because the average male pitch period is longer than for females. Whether this difference would be enhanced or attenuated in much larger window sizes (e.g. 64 ms) is not known,

although the accepted rule of thumb in the speech community is that the best window size for males and females differs.

Phoneme Category:

Vowels

BEST OVERALL WINDOW FOR VOWELS: 8 and 16 ms G, H, T

Across exemplars, the 16 ms window gave a slight advantage for males, and the 8 ms window gave a slight advantage for female speakers.

Fricatives

BEST OVERALL WINDOW FOR FRICATIVES: 8 or 16 G, H, T

The criteria for choosing the best window was: strength of frequency information, frequency resolution at onset of phoneme, and presence/strength of F0 for voiced fricatives. This latter criteria was somewhat problematic. /F/ (which is theoretically voiceless) often showed stronger evidence of F0 than some very weak exemplars of /V/. Further, the presence of F0 varied considerably with window size and shape, so it was often difficult to determine for certain whether an apparent F0 was actually present, or was merely a windowing artifact.

There appeared to be a tradeoff between high and low frequency information across window size. The 4 ms window was more apt to clearly show high frequency energy, while the 16 ms window was most likely to pick up the F0, usually at the expense of high frequency information. Since time information is relatively less important in fricatives (than in, say, stops) the window with

the best frequency resolution was chosen as the best, as long as the onset/offset of frication was clearly identifiable.

Stops

OVERALL BEST WINDOW FOR STOPS: 4 or 8 ms G, H, T

The criteria for best window was frequency resolution and time resolution at the CV transition. Since evidence of voicing was difficult to determine for stops, it was not included in the judgment criteria. For stops there was a trade-off between time and frequency resolution with the 4 ms window giving slightly better time, and the 8 ms window slightly better frequency resolution. Since the differences were not large, it was generally impossible to choose one over the other - both did an acceptable job in both parameters. The 16 ms windows, on the other hand, had very poor time resolution with no useful gain in frequency resolution.

Summary and Conclusion

The results of this study confirm the hypothesis that different window shapes and sizes are best suited to the analysis of different types of phonemes. For stops, the 4 ms Gaussian, Hanning and Triangular windows gave the best results across exemplars while the Rectangular, the Exponential, and all 16 ms windows gave quite poor results. For fricatives, the 8 and 16 ms Gaussian and Triangular windows gave the best results. The Hanning window was poor for fricatives, because it often failed to give F0 information.

For vowels, the 8 or 16 ms Hanning windows were best for male speakers, and the 8 ms H window was best for female speakers. The 8 and 16 ms Gaussian and Triangular windows also gave good results, for vowels, and all 4 ms windows gave poor results. A few spectrograms created with much larger (64 ms) windows, however, suggested that until these much larger windows are investigated, no final conclusions can be made about the best window size for vowels.

REFERENCES

- Graff, D. E., and W. I. Lundgren. (1987). Speaker independent connected speech. Final Technical Report RADC-TR-87-85. Rome Air Development Center, Griffiss Air Force Base, New York.
- Harris, F. J. (1978). On the use of windows for harmonic analysis with the Discrete Fourier Transform. Proceedings of the IEE, 66, 51-83.
- Kent, R. D. (1983). The segmental organization of speech. In P. F. McNeilage (Ed.) The production of speech. New York: Springer-Verlag.
- Lineback, J. R. (1984). Recognition system processes speech the way the ear does. Electronics, 57, p. 45-6.
- Nathan, K.S., and H. F. Silverman. (1990). High-resolution characterization of formants in vowel-consonant transitions. ICASSP, 1, p. 237-240.
- Rhody, H. E., R. A. Houde, C. V. Parkins and S. Dianat. (1985). Speech analysis based on a model of the auditory system. Unpublished manuscript. Rome Laboratories, Griffiss AFB, New York.

Appendix A: SPEECH DATA USED

| TIMIT SENTENCE NUMBER | TEXT OF SENTENCE / PHONEMES ANALYZED |
|-----------------------------|---|
| SA1 | She had your <u>dark</u> <u>suit</u> in <u>greasy</u> wash water all year. /d/ /s/ /aa/ /ux/ /iy/ /k/ |
| SA2 | Don't <u>ask</u> <u>me</u> to <u>carry</u> an oily <u>rag</u> like that. /ow/ /ae/ /iy/ /k/ /ae/ /s/ /t/ |
| SX14 | <u>Before</u> Thursday's exam, <u>review</u> <u>every</u> <u>formula</u> . /b/ /v/ /v/ /f/ /f/ |
| SX23 | <u>Those</u> musicians harmonize <u>marvelously</u> . /ow/ /z/ /z/ /aa/ |
| SX35 | Help <u>celebrate</u> your brother's success. /b/ |
| SX56 | <u>Academic</u> <u>aptitude</u> <u>guarantees</u> your diploma. /d/ /ux/ /g/ /t/ |
| SX146 | Cyclical <u>programs</u> never <u>compile</u> . /p/ /g/ /p/ |

APPENDIX B: SUMMARY OF BEST WINDOWS BY PHONEME EXEMPLAR

| VOWEL | CONTEXT | MALE SPEAKER | FEMALE SPEAKER |
|-------|-----------|--------------------------|--------------------------|
| AA | DARK | 16 ms G H T | 8 ms G H T (ALL POOR) |
| | MARVELOUS | ALL POOR | 16 ms G H T |
| AE | ASK | 16 ms G H T (8 ms G H T) | 4,8 ms G H T (ALL POOR) |
| | RAG | 8 ms H T (POOR) | 4 ms G H T |
| IY | ME | 16 m H (8 G H T, 15 G T) | 16 m H (8 G H T, 16 G T) |
| | GREASY | 8 m H (8 G T, 16 G H T) | 16 m H (8 G H T, 16 G T) |
| OW | DONT | 8 ms G H T | 8 ms G H T |
| | THOSE | 4 ms G H T | 8 ms G H T |
| UX | APTITUDE | 16 ms H (8 ms H) | 8 ms H R |
| | SUIT | 16 ms H | 4 ms T, 8 ms G H T |

| FRICA- TIVE | CONTEXT | MALE SPEAKER | FEMALE SPEAKER |
|----------------|-----------|-------------------|-------------------|
| F | BEFORE | 8,16 ms T G | 8, 16 ms G T |
| | FORMULA | 8,16 ms T G | 8, 16 ms G T |
| V | EVERY | 8 ms E G R T | 8 ms G T |
| | REVIEW | 8 ms E G T | 8, 16 ms E G T |
| S | ASK | 4, 8, 16 ms G H T | 4, 8, 16 ms G H T |
| | SUIT | 4, 8, 16 ms G H T | 4, 8, 16 ms G H T |
| Z | HARMONIZE | 16 ms G T | 16 ms G T |
| | MUSICIAN | 16 ms G T | 16 ms G T |

| STOP SPEAKER | CONTEXT | MALE SPEAKER | FEMALE |
|-----------------|------------|-------------------------|--------------------------|
| P | COMPILE | 8 ms G H (4, 8 ms T) | 8 m H (4 m E, 4,8 G R T) |
| | PROGRAM | 4, 8 ms ALL WINDOWS | 4, 8 ms G H T |
| B | BEFORE | 4 ms H (4 ms G T) | 8 ms H (4 m T, 8 m E G) |
| | CELEBRATE | 4 ms G (4 ms E H R T) | UNREADABLE |
| D | ACADEMIC | 4 ms G H T (8 ms G H T) | 4 ms G H T (8 ms G H T) |
| | DARK | 4 ms G H T | 4 ms G H T (8 ms G H T) |
| T | APTITUDE | 4 ms G H T | 4 ms G H T |
| | ME#TO | 4 ms G H T | 4 ms G H T |
| G | GUARANTEES | 4 ms G H T | 4 ms G H T |
| | PROGRAMS | 4 ms ALL WINDOWS | 4 ms H (8 ms E G T) |
| K | CARRY | 4, 8 ms E G H T | 4 ms E G H T |
| | DARK | 8 ms R (4 ms R G H T) | 8 ms E (4 ms E G H T) |

RTS Prototyper's Workbench:

A Tool for Rapid Prototyping Provably Correct Real-Time Systems

by

James F. Peters III

*Syracuse University
School of Computer & Information Science
4-116 CST, Syracuse, NY 13244-4100
and*

*Computer Science Department
University of Arkansas--Fayetteville
232 Science Engineering Building
Fayetteville, AR 72701*

email: peters@top.cis.syr.edu

A Research Report

submitted to

**Research & Development Laboratories
Culver City, CA 90230-6608
U.S.A.**

1991

RTS Prototyper's Workbench:

A Tool for Rapid Prototyping Provably Correct Real-Time Systems*

James F. Peters III

*Syracuse University
School of Computer & Information Science
4-116 CST, Syracuse, NY 13244-4100*

*and
Computer Science Department
University of Arkansas--Fayetteville
232 Science Engineering Building
Fayetteville, AR 72701*

email: peters@top.cis.syr.edu

Abstract

This paper reports work on a prototyper's workbench, which is a software tool for rapid prototyping provably correct real-time systems. A real-time system (rts) is conceptualized as plant || controller (read "cooperation between hardware and controller software used to control the operation of the hardware"). The principal aim of this workbench is to provide a formal basis for prototyping provably correct software for a real-time system. The workbench has three layers: visualization, specifier's assistant, and experimentation. It supports reusability of design concepts with libraries of existing visual representations and temporal specifications. The workbench can also be used to model the environment for an rts, which makes it possible to test the behavior of a prototype. The focus of this report is on the specification layer of the workbench. Specifications are written in Ada/TL_{rt}, which is a language for specification of the behavior of real-time systems. It merges concepts of the specification part of Ada, VDM specification of packages, and temporal logic specification of task behavior. Ada/TL_{rt} specifications utilize explicit clock temporal logic, which facilitates formulation of hard, real-time constraints with respect to an external clock. These specifications prescribe constraints on state transitions, properties of behaviors of individual tasks and properties of interacting tasks. A proof of a system specification consists of showing that the system property holds over all possible interleavings of the task behaviors. An example real-time system (traffic light controller) specified in Ada/TL_{rt} along with a partial proof of its specification, is presented.

Index Terms -- Ada, hard real-time, prototyping, program specification, real-time systems, temporal logic, visualization.

* Research supported in part by the Research & Development Laboratories, Culver City, CA 90230-6606 USA.

1. Introduction

Considerable work has been done on describing the behavior of hard, real-time systems [Alu90; Cro88; Har90; Han90a, Han90b; Hen91; Kop88, Kop89a, Kop89b; Ost89, Ost90; Pet90a, Pet90b, Pet91a, Pet91b, Pet91c]. A *hard, real-time system* (rts) is a computer system where the validity of results produced by the rts depend on both logical correctness and timeliness [Kop89a]. An rts consists of two main parts: controller and plant. The controller is a computer which processes input from the environment as well as plant and supplies control information to the plant (hardware). The critical features of a controller are synchronization (rendezvous), concurrency (concurrent behaviors of communicating processes), responsiveness (behavior which adheres to timing constraints), determinism (behavioral transitions which satisfy enabling conditions), and non-determinism (interleaving of observed behaviors of concurrent processes). This paper reports work on a Prototyper's Workbench, which is a software tool for rapid prototyping provably correct real-time systems. The principal aim of this workbench is to provide a set of tools for developing provably correct software for an rts. The workbench has three layers: visualization, specifier's assistant, and experimentation.

The focus of this paper is the specification language Ada/TL_{rt} and its proof system in developing provably correct rts software with the Specifier's Assistant of the Prototyper's workbench. Ada/TL_{rt} extends Ada/TL [Han 90a, Han 90b, Pet 90a, Pet 90b], which is an Ada-based specification language which utilizes standard, temporal logic. Standard Ada specification of tasks consists of declarations of tasks and related data types [Dod83]. Ada/TL consists of (i) VDM style specification of package and task operations [Jac85, Jon90], (ii) local temporal specifications about the behavior of each task written in temporal logic from [MP 81, MP 83], and (iii) and global system properties that specify required constraints on the interaction between tasks using branching time operators of UB logic [Ben81]. Ada/TL_{rt} replaces the standard, temporal logic in Ada/TL with explicit clock temporal logic [Ost 89, Har 90, Hen 91]. As a result, Ada/TL_{rt} permits the specification of hard, real-time constraints on the behavior of communicating tasks.

Pnueli and others [Pnu 79, MP 81, MP 83, Lam 83] advocated use of temporal logic for specification of the behavior of concurrent programs and the development of temporal proof systems. Barringer and Mearns [Bar 82] and Pnueli and DeRoever [Pnu 82] presented proof systems for parts of Ada, which were concerned with verification of

Ada code, not with verification of the specifications. Clark et al. [Cla 85] illustrated a verification method for assertions for finite state systems, which are similar in concept to the computation graphs described in this paper. Young [You 85] also identifies the need to verify specification separately from the work of verifying code.

The proof system for Ada/TL_{rt} consists of (i) notations for representing states within a computation graph, (ii) inference rules for justifying transformations between states, with each rule corresponding to objects and temporal operators in Ada/TL_{rt}, (iii) axioms about progress and fairness in the tasking computation, (iv) a constructive proof of properties satisfied over states of the computation. Proof of an Ada/TL_{rt} specification consists of showing that the system specification is consistent with all local task specifications. Task properties are taken as premises in the proof of the system property. Once the system property has been verified, it does not need to be verified again for the program implementation. The task specifications are intended as the basis for designing the corresponding task bodies. Proving the tasking specification will reduce verification of the corresponding Ada code to showing that each task satisfies its own local specification. The aim of the underlying research is to provide a formal foundation and a methodology for constructing proof tableaux for tasking system properties. In practice, a proof tableau would be analyzed using an automated theorem prover such as a Nuprl [Con 86], which is a part of the Specifier's Assistant in the Prototyper's Workbench. The extraction of a real-time system program from the proof of its specification is discussed in [Pet 91b].

An overview of the structure of the Workbench is given in Section 2. The idea of visualizing the behavior of communicating processes with temporal input/output automata is presented in Section 3. The notion of a Specifier's Assistant embodied in the Workbench is given in Section 4. The form of explicit clock temporal logic (called TL_{rt}) used by the Workbench is presented in Section 5. This form of temporal logic relies on rigid and flexible variables to define the semantics of timing constraints. The usage of these variables is explained in Section 6. TL_{rt} provides a basis for a duration predicate $\text{delay}(k)$, which can be used to specify either a lower bound (Section 6.2) or an upper bound (Section 6.3) on a system action. A sample Ada/TL_{rt} specification for a traffic light control system is given in Section 7. Computation graphs derived from the proof of an Ada/TL_{rt} specification are presented in Section 8. The axioms, rules, metalanguage (called Temporal Observer's Language), proof structure, and example proof of a specification are given in Sections 9 and 10. A brief discussion of the experimentation layer of the Workbench is presented in Section 11.

2. Structure of the Workbench.

The workbench consists of a set of layers, which are non-hierarchical:

- **Visualization Layer.**

The visualization layer is used to model the timed-behavior of a real-time system. It has an optional natural language interface which can be used to construct timed, predicate i/o automaton visualizations of behavior derived from requirements expressed in ordinary language. Timed automata can either be constructed by hand or in conjunction with a library of existing automata.

- **Specifier's Assistant Layer.**

The specifier's assistant layer provides an environment for writing what we call constructive, temporal specifications of a prototype system. The behavior of plant and controller are expressed as deterministic, timed predicate i/o automata. A provision is made for specifying constraints on the interaction between components of the controller, which is usually a distributed computer control system with components communicating with each other over a local area network. Constraints on the behavior of an rts are formulated with the help of a knowledge base, which contains knowledge about particular applications. The byproduct of this layer is a prototype program, which is constructed from the proof of the specification.

- **Experimentation Layer.**

The experimentation layer provides an interactive environment for the testing of prototypes with the help of simulators.

3. Visualization of System Behavior.

The Visualization Layer of the workbench provides a structured notation for representing the behavior of a real-time system. This notion makes it possible to construct finite, directed, labelled graphs called timed, predicate i/o automata (TA_{i/o}s). This layer has a library of timed automata, which can be composed with each other and/or with designer-supplied automata to form a prototype system. That is, when automata are composed with each other, they are "connected" together via input/output channels which permit communication between the automata. The behavior of both the hardware and the controller can be visualized. Each constructed automaton represents what is known as an agent. An *agent* is that part of a real-time system which has its own identity which persists over time and its own externally observable behavior. The nodes in these graphs represent agent states (each node is labelled with a predicate which specifies the activity associated with the agent state as well as hard, real-time constraints on the duration of a state activity). The arcs of these graphs represent transitions between states (each arc is labelled with an enabling condition, which must be satisfied before the transition along the labelled arc can occur). The labelling of these graphs provides an indication of the actions and hard, real-time constraints (read "required responsiveness") of an RTS component. The formulation of the predicates used to label the nodes and arcs of a graph is done with the help of a natural language interface. The role of TA_{i/o}s in constructing rts from system specifications is given in [Pet 91b].

4 Specifier's Assistant

The Specifier's Assistant Layer of the Workbench provides a formal basis for constructing concurrent programs from proofs of their specifications. Again, this layer provides an agent-oriented specification technique (processes are viewed abstractly as agents). The behavior of individual agents is represented assertionally with explicit clock temporal logic in the specifier's layer. As in the visualization layer, there is a library of temporal specification templates. These templates are temporal formulas which can be used to facilitate the development of the specification of a new system. Each library template has a corresponding graph (if designers choose to, they can "see" the behavior of the system they are specifying assertionally). In addition, designer-supplied temporal formulas can be used to construct a corresponding timed i/o

automaton. Underlying this specification technique is an interest in formulating reasonable, hard, real-time constraints on the behavior of individual agents as well as predicting system behavior.

The specifier's assistant utilizes a Temporal Observer Language TOL, which is a metalanguage used to make observations about the satisfaction of temporal formulas. TOL provides a formal basis for model checking and a notation useful in proving the satisfaction of specifications in an automatable proof tableau. So far in this research, a form of distributed Ada has been used as the target language [Wen90, Pet 91a].

5 Explicit Clock Temporal Logic

The behavior of a real-time system can be specified with Real-Time Temporal Logic (RTTL) given in [Ost 89, Har 90, Hen 91]. RTTL is an explicit clock logic which uses data variables to reference an external clock in assertions. When temporal logic is applied to the study of processes, the formulas of temporal logic are interpreted as predicates over sequences of process states [Alp 86]. Each state occurs at some instant in time in which the values of process variables can be inspected. During a succession of states, changing values of state variables may entail changing truth values of predicates about state variables. Temporal formulas can be used to enumerate state transitions (transformations of one state into a new state) in a behavior as well as the order in which transitions are made.

RTTL provides a concise means of prescribing a property of a behavior of an rts controller or plant; such prescriptions are assertional. This form of temporal logic is essentially the same as the original temporal logic introduced by Manna and Pnueli [MP 81, MP 83] with the addition of data variables such as T (for timing constraints) suggested by [Har 90, Hen 91]. Except for some additional derived temporal operators taken from [Pet 90a, Pet 91b], the temporal logic used in this article is the same as RTTL. For simplicity, we limit the presentation of RTTL to a discussion of the U (until) and temporal operators derived from U . We also introduce the derived temporal operators before, $\Diamond \omega$ (infinitely often), and $\text{seq}(p_1, p_2, p_3, \dots, p_n)$ (a temporally quantified sequence of state predicates where p_1 holds before p_2 , which holds before p_3 , ..., before p_n). For the subset of RTTL (named TL_{rt}) we have chosen, the temporal language TL_{rt} is defined as follows:

Alphabet

- A denumerable set of variables: x, \dots

- A denumerable set of n-ary functions: f, g, \dots
- A denumerable set of n-ary predicate symbols: p, q, \dots
- Symbols \neg , or, \forall , $(,), U$

Well-formed formulas of TL_T have the following syntax:

- Every atomic formula is a formula.
- If x is a variable and A is formula, then $\forall x A$ is a formula.
- If A and B are formulas, then $\neg A$, $(A \text{ or } B)$, $(A U B)$ are formulas.

5.1 Semantics of Temporal Operators.

The \neg (not), or, and \forall (all) symbols have the usual semantics. In addition, the implication symbol \Rightarrow (i.e., $p \Rightarrow q \equiv \neg p \text{ or } q$) is used. In defining the following semantics, the notation

$$(q_0, \dots, q_x) \models p \text{ for } x \geq 0$$

asserts that each of the states in the sequence (q_0, \dots, q_x) satisfy predicate p . In what follows, let q_0 represent the current state in a behavior. Let p, q be first-order predicates. The semantics of U as well as the operators derived from U are as follows:

$$p U q \equiv \exists k, x: 0 \leq x \leq k: (q_0, \dots, q_x) \models p \text{ and } q_k \models q$$

$$p \text{ before } q \equiv \exists k: 1 \leq k: q_0 \models p \text{ and } (q_1, \dots, q_k) \models p U q$$

$$\Diamond p \equiv \text{true } U p$$

$$q_k \models \text{seq}(p) \equiv q_k \models p$$

$$\text{seq}(p_1, (\text{seq}(p_2, \dots, p_n))) \equiv p_1 \text{ before } \text{seq}(p_2, (\text{seq}(p_3, \dots, p_n)))$$

$$\Diamond \omega p \equiv \text{seq}(p, \Diamond \omega p)$$

The predicate ' $p U q$ ' asserts that the predicate q eventually holds (either in the current or in some future state) and that the predicate p holds in the current state and in each of the states *until* the state when q holds. By contrast, ' p before q ' asserts that p is guaranteed to hold initially and sometime later q will hold. For this reason, *before* is called a precedence operator [Krö 85]. These powerful temporal operators provide the

basis for the semantics of the remaining operators in the above list. TL_{rt} is used to specify timing constraints on controller actions.

Branching time quantifiers are defined in a similar fashion, but over composite states of a system. A system state S_i is a tuple $\langle G_i, s_{ai}, s_{bi}, s_{ci}, \dots \rangle$ where G_i is a record of all global data items, and $s_{ai}, s_{bi}, s_{ci}, \dots$ are states of the component tasks a, b , and c of the system. A system path is denoted as $H = (S_0, S_1, S_2, \dots, S_i, \dots)$. TL_{rt} utilizes the branching-time quantifier $\forall \square$. An assertion of the form $\forall \square p$ asserts that p holds in all system states of all system paths:

$$\forall \square p \equiv (\forall i, j \text{ in Nats, } S_i \text{ in } H_j: \text{sat}(S_i \mid H_j, p))$$

6. External Clocks and Timed Behaviors

Explicit clock logic relies on the use of rigid and flexible variables. A *rigid variable* is a variable r which can be assigned a value in a particular rts state and r retains its value across state changes. By contrast, a *flexible variable* value can change with state changes. For example, the rigid variable T records the Clock value. We assume that the value of T can be changed when needed (this is analogous to resetting the external clock in a timed Büchi Automaton [Alu 91]). The flexible variable *time* gives the value of Clock in the current state. Clock readings are non-negative, real numbers. Each time an event occurs, a reading of Clock is associated with that event.

6.1 Semantics of Delay

Responsiveness of a system is measured in terms of actual values of delays. The duration predicate $\text{delay}(k)$ asserts that the external clock is allowed to run for k ticks before a timeout occurs. $\text{Delay}(k)$ can be used to specify a lower bound on the number of ticks before an action is performed; $\text{delay}(k)$ can also be used to specify an upper bound on the duration of an action. A similar technique for specifying timing constraints on actions is used by [Hand 88].

6.2 Lower Time Bound on System Actions

We can express a lower bound on the number of ticks before a system action begins. If we let ACT be the action to be performed in state q . We can express the fact that we let the external clock run for k ticks before performing ACT by writing informally

"delay(k) before ACT." To see this, let T record the time in state q . Assume action ACT is performed in state q . Written by itself, "ACT" is shorthand for the assertion "the action ACT is performed." Let $\text{sat}(q \mid (q'), P)$ mean that predicate P is satisfied in state q of the state sequence (q, q') , and $\text{sat}(q', Q)$ mean that predicate Q is satisfied in state q' . The double turnstile \models reads "forces" or "satisfies." Then satisfaction of "delay(k) before ACT" over a state sequence (q, q') is expressed in clausal form as follows:

$\text{sat}(q \mid (q'), \text{delay}(k) \text{ before ACT}) :-$
 $q \models \text{delay}(k) \text{ and } T \leq \text{time} < T + k,$
 $q' \models \text{ACT and time} = T + k.$

This says that the duration predicate is satisfied in state q and k ticks later the predicate ACT is satisfied in state q' .

6.3 Upper Bound on Actions

We can also express an upper bound on the number of ticks on the duration of a system action using $\text{delay}(k)$. This is expressed rather simply by writing "ACT; $\text{delay}(k)$," which asserts that ACT cannot be continuously enabled for more than k ticks of the external clock. The predicate *timeout* is an enabling condition, which evaluates to true at the k th tick of the clock (i.e., an action which must be performed within k ticks times out, and a transition to the next state occurs). The meaning of this upper bound constraint can be explained concisely by using the satisfaction clause $\text{sat}(q, P)$. Then the upper bound timing constraint can be defined as follows:

$\text{sat}(q, \text{ACT; delay}(k)) :-$
 $q \models \text{ACT},$
 $q \models \text{time} < T + k; \quad /* ; \text{ reads "or" } */$
 $q \models \text{time} = T + k \text{ and } \text{timeout}.$

7. Ada/TL_{rt} Specifications

This section illustrates Ada/TL_{rt} notation using an example WalkLightSystem. First, an Ada specification of the system is given in Figure 1a.

```

package WalkLightSystem is
  DirectionType is (NS, EW);
  ColorType is (Red, Green);
  Light : array(DirectionType) of ColorType ;

  task type WalkerTask is
    -- cycle SelectDirection,
    --if SeeGreen then Walk
    --else ( LightController.PushButton(Direction) and wait, Walk )
    --endif;
    -- end cycle
  end WalkerType;

  task LightController is
    entry PushButton(DirectionType);
    --cycle alternately for EW and NS,
    --delay(MinWalkTime);
    --wait for PushButton or delay(WaitTime) ;
    --ChangeLight; clear PushButton queue
    -- end cycle;
  end LightController;

  Walker: array ( positive range <> ) of WalkerType;
  --System property: all concurrent Walkers have the same direction;

end WalkLightSystem;

```

Fig.1a WalkLightSystem Ada Specification

The WalkLightSystem allows an arbitrary number of Walker tasks which represent walkers crossing through a controlled intersection. Walkers select either a NS or EW direction. If the light in the selected direction is Red, then the Walker pushes a button connected to the LightController and waits for the Light to turn Green. The LightController regulates the flow of traffic through the intersection. The Light changes periodically after a minimum period MinWalkTime if a button is pressed or after a maximum period of (WaitTime ticks of an external clock). The specification is expanded in Ada/TL_R form in Figure 1b.

```

package WalkLightSystem is
  DirectionType is (NS, EW);
  ColorType is (Red, Green);
  Light:array(DirectionType) of ColorType := (Red, Red);

  task type WalkerType is
    Direction: DirectionType ;
    procedure Walk;
    procedure SelectDirection is

```



```

out Direction = NS or Direction = EW;
def SeeGreen = (Light(Direction) = Green);
Property
  ◇ω seq(
    SelectDirection,
    ( SeeGreen and Walk
      or
      ¬ SeeGreen and seq(LightController.PushButton(Direction), Walk)
    )
  );
end WalkerType;

task LightController is
  MinWalkTime, WaitTime, RedTime:
    constant DURATION pending;
  RedDirection: DirectionType := NS;
  GreenDirection: DirectionType := EW;
  procedure BothRed is
    out: seq (Light = (Red, Red), delay( RedTime)) ;
  def Old = if RedDirection=NS then EW else NS endif;
  procedure ChangeLights is
    out: seq (BothRed,
      if RedDirection=NS
      then Light=(Green, Red) and RedDirection = EW
      else Light=(Red, Green) and RedDirection = NS
      endif);
  procedure FlushQueue(Direction: IN DirectionType) is
    out: |PushButton(Direction)'COUNT| = 0
  procedure IsClear(GreenDirection) is
    --code for LightController to determine if intersection is clear in green direction.
  entry PushButton(Direction: IN DirectionType);
  property
    ◇ω seq ( delay( MinWalkTime), --lower bound on PushB.
      when IsClear(GreenDirection) =>
        PushButton(RedDirection); delay(WaitTime), --upper bd on PushB.
        --response is WaitTime

      when ¬ timeout =>
        seq(ChangeLights,
          FlushQueue(RedDirection) )
    );
  end LightController;
Walker: array ( positive range <> ) of WalkerType;
SystemProperty
  ∀□ (∀ i, j: i ≠ j: Walker(i).Direction ≠ Walker(j).Direction
    ⇒ ¬ (Walker(i): Walk and Walker(j): Walk))
end WalkLightSystem;

```

Fig.1b WalkLightSystem Ada/TL Specification

Each task Walker selects a direction and if the appropriate Light is not Green, the Walker calls to rendezvous with the appropriate PushButton of the LightController. The LightController cycles continually. For each cycle, it delays the MinWalkTime, and then if the green direction of the intersection is clear and there is a client pushing the button within WaitTime ticks of the external clock, then the change lights sequence of the LightController begins. That is, the rigid variable T stores the reading of the external clock after MinWalkTime ticks, and a change of the lights must be requested within $T + \text{WaitTime}$ ticks of the external clock. If the LightController has entries on the PushButton queue for the Red Light, then all of those entries are cleared. Ada/TL_{rt} utilizes the task attribute STATE, which is a record of all information related to states of execution of tasks. In addition to the two task specifications, the package specification contains a **SystemProperty** which specifies task interactions not constrained by the task specifications. The SystemProperty asserts that for all execution sequences generated by the WalkLightSystem, Walkers wishing to walk in different Directions do not Walk at the same time.

Task specifications are augmented with definition of static variables like Direction, and operations (procedures or functions) like Walk, which are referenced as part of the task specification. We refer to these items as externally observable, in that they are used in description of the task specification but they are not accessible by other tasks, nor are they necessarily part of the task implementation. The observable items may become items in the task body or they may correspond to auxiliary variables that are used only in the proof of the system behavior. Observable operations and entry specifications are annotated with non-temporal "in" and "out" assertions, as in VDM specifications. The last component of each task specification is the property assertion which specifies the linear-time temporal behavior of the subject task, including interaction with other tasks and access to either global items or its own observable items.

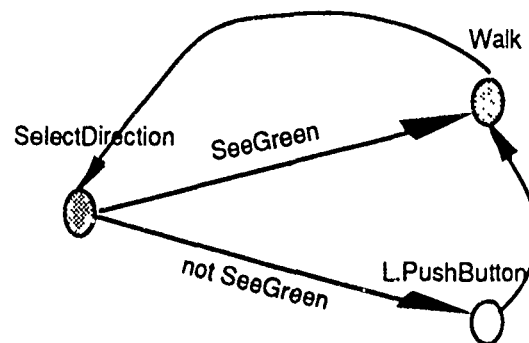
The predicate "Walk" is interpreted as the control predicate "at(Walk)". The predicate "LightController.PushButton(Direction)" is interpreted as the control predicate "at(LightController.PushButton(Direction))" which means that the subject task calls the entry PushButton(Direction) in task LightController with the previously instantiated value of Direction. The **SystemProperty** constrains interactions of multiple tasks. We use the notation TaskName:Op to indicate task TaskName is performing operation Op. The " $\forall \square$ " operator quantifies the assertion following it in

branching time and reads "for all possible execution sequences generated by the program".

In summary, non-temporal predicates express constraints about parameters and observable variables. They constrain instances of "in" parameters and they instantiate values of "out" parameters. Control predicates express events of execution of entry procedures and observable operations. They are interpreted as using the "at" predicate [Lam 83], although the "at" is not written. Temporal predicates are composed using temporal quantifiers applied to either non-temporal or control predicates. Temporal predicates express sequence, repetition, or eventuality of predicate conditions.

8. Computation Graphs

In concept, possible behaviors of every specified task can be represented as a computation tree. The computation tree is a rooted, finitely branching tree with nodes corresponding to predicates about single behavior states and arcs labeled with predicates about branching conditions. Any single path down through the tree defines a state sequence h . The tree is rooted at s_0 of h . More practically, the behaviors represented in a computation tree can be folded into a finite state computation graph. The computation graph for a Walker task is shown in Figure 2.



Legend: L = LightController

Fig. 2 Transition graph for Walker task.

The computation graph is an I/O automaton in that certain state predicates correspond to interactions with computation graphs of other tasks. The behavior of a tasking system

can be represented as a system of interacting computation graphs. For the WalkLightSystem, we can draw the computation graph shown in Figure 3.

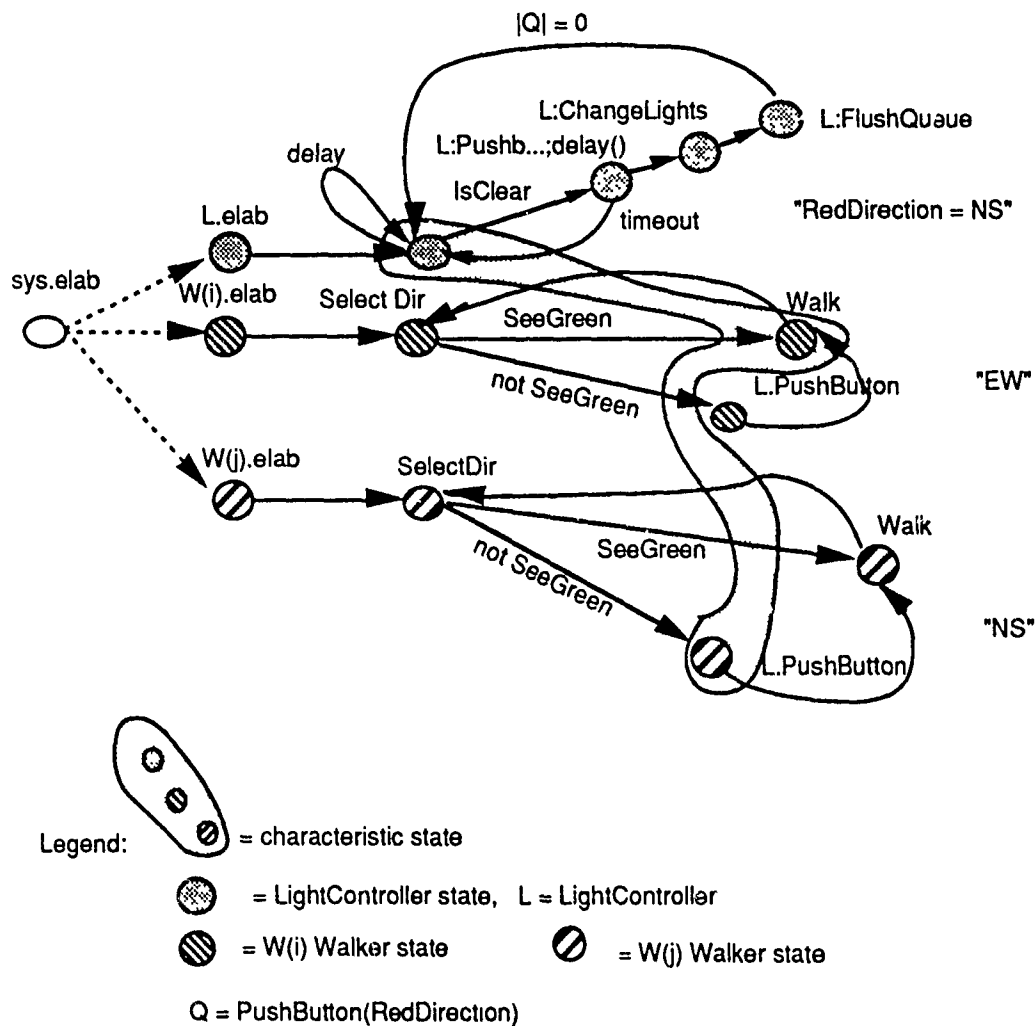


Fig. 3 Computation graph for WalkLight system.

The predicate for any system state consists of the conjunction of predicates for the states of all the component tasks. Even simple tasking systems generate many possible behaviors because of possible non-deterministic interleaving of individual task behaviors. To analyze system behavior, we identify characteristic states. A *characteristic state* is a system state that occurs repeatedly in the system behavior. For the WalkLightSystem one characteristic state is described by the predicate:

AND(∇ L:PushButton and RedDirection=NS,

- ▼ W(1):Walk and Direction = EW,
- ▼ W(2):L.PushButton(NS)).

The ▼ symbol, read as "at or before," is explained later. In proving that a system property is satisfied by the local properties of system tasks, we first identify characteristic states. We show the behavior of processes eventually leads to a characteristic state. Then we show that the desired system property holds over each path from the initial state to the characteristic state and from the characteristic state back to the characteristic state.

9. Proof System

The proof system is comprised of inference rules for the constructs used in Ada/TL_{RT} and the actual proof method. This section presents a minimal set of inference rules and the general proof structure. The essence of the proof technique is to show evaluation of the system property over all system paths that can be composed from task paths determined by the task properties.

9.1 Notation

To verify system properties, we need notation for associating predicates with task states. Such notation is part of what is known as an observer language for task behavior. We use the marker operator ▼, read as "before, or at". The ▼ marker is not a temporal operator, but it is used to denote the state of a task in terms of its specification. Interpretation of marked predicates is explained next. There are two interpretations of marked predicates, first as a trace of possible task behaviors, and second, as a trace of task states in a system proof. We show first the interpretation of marked predicates for tracing behaviors. Let h (a state history) be a sequence of task states, and P be a specification. We say $\text{sat}(h, \nabla P)$, if the history h satisfies the predicate P . At any point in the observation of system behavior, an individual task can be at some point "between" observable states; let such an internal state be represented by τ . Any general predicate is composed of predicates about single states using the temporal operators. A predicate about a single state is called atomic. An atomic predicate has the form

$$p = p_E \text{ and } p_C$$

where p_s is called a state predicate and it constrains values of all the task local variables, p_c is called a control predicate and it constrains the single state operation. For example, the predicate (Direction=EW and Walk) is atomic. We denote atomic predicates with the metapredicate "isatomic". The relation sat has the following properties:

$\text{sat}(\text{null}, \nabla \text{TRUE})$.
 $\text{sat}((\tau, h), \nabla P)$ if $\text{sat}(h, \nabla P)$.
 $\text{sat}(s, \nabla p)$ if $\text{isatomic}(p)$ and $p(s) = \text{TRUE}$.
 $\text{sat}((h_1, h_2), \text{seq}(p_1, \dots, \nabla p_i \dots p_n))$
 if $\text{sat}(h_1, \nabla p_i)$ and $\text{sat}(h_2, \text{seq}(p_1, \dots, \nabla p_{i+1}, \dots, p_n))$.

We will use the following shortened notation to simplify marked expressions. If the specification for a task is of the form $P = \text{seq}(p_1, \dots, p_n)$, then we write ∇p_i to represent the expression $\text{seq}(p_1, \dots, \nabla p_i, \dots, p_n)$ and we write $p_i \nabla$ to indicate completion of the state p_i . The following equivalence rules are given without proof:

$\nabla (p \text{ or } q) \equiv \nabla p \text{ or } \nabla q$
 $\nabla \Diamond P \equiv \nabla P$
 $\nabla \Diamond \omega P \equiv \text{seq}(\nabla P, \Diamond \omega P)$
 $\nabla \text{seq}(p_1, \dots, p_n) \equiv \text{seq}(\nabla p_1, \dots, p_n)$
 $\text{seq}(p_1, \dots, p_i \nabla, \dots, p_n) \equiv \text{seq}(p_1, \dots, \nabla p_{i+1}, \dots, p_n)$
 $\text{seq}(p_1, \dots, p_n \nabla) \equiv \text{seq}(p_1, \dots, p_n) \nabla$

We show second the interpretation of marked predicates as states within a system proof. In developing a proof of system behavior, we trace states of each component task of the system, using marked predicates to indicate states. A proof does not argue that a particular path satisfies the state predicates. Rather, each step is a justification that all possible paths from previous states satisfy the state predicate. Justification for each step is based upon axioms and transformation rules given in the following sections. In proofs, the notation ∇p is interpreted to indicate that the task is in a state or will next be in a state which satisfies p (the task could be in a τ state). The notation p (without the marker) indicates that the task is in a state which satisfies p . The notation

$p \blacktriangledown$ indicates that the task has completed the operation associated with the control predicate part of p .

9.2 Axioms and Rules

The proof system for reasoning about the interactions of communicating agents, carries over to Ada/TL_{rt} specifications of tasking systems. In the context of Ada, these rules have interpretations summarized in Table 1. Inference rules in this section and the next section will be used to verify each state in the proof. The axioms in this section are rules which are assumed or required. Rules of the form " q infer p " assert that if condition q holds for a preceding state(s), then condition p can be concluded to hold in the subject state. Rules of the form

q
infer
(i) p ,
(ii) r

assert that a state that satisfies q can lead to two different states that satisfy p and r .

Table 1 Interpretation of Proof System in Ada Context

Note: $p = p_{state}$ and $p_{binding}$

C, S, E = client task, server task, and task entry, respectively.

$inx(C, S, E) == (C:S.E \text{ and } S:E \text{ and } S:STATE.Q(E).CallerId = C)$

$P_i = seq(p_{i1}, \dots, p_{im})$ and $isatomic(p_{i1})$

Axiom/Rule

Interpretation_i (in Ada/TL_{rt})

Progress/Fairness Axioms:

P.1a $\blacktriangledown p$ infer p

p represents either an Ada procedure call (intra-op) or entry-call or accept call (inter-op).

P.1b p infer $p \blacktriangledown$

p (same as in P.1a)

- | | |
|---|---|
| <p>P.2 $\text{inx}(C, S, E) \nabla$ infer $C.\text{STATE.Beh.params'OUT'new}$ $= C.\text{STATE.Beh.params'OUT'old}$ $+ \text{callrec}(C, \text{params'OUT})$ and $C.S.E \nabla$ and $S:E \nabla$</p> | <p>Ada rendezvous</p> <p>params'OUT received by Client from Server.</p> <p>After entry call, after accept</p> |
| <p>F.1 $\nabla \Diamond^\omega (S:e_1 \text{ or } \dots \text{ or } S:e_n)$ infer $\nabla S:e_i$</p> | <p>(fairness) ea. entry e_i is eventually selected.</p> |
| <p>B.1 $\nabla \text{OR}(P_1, \dots, P_n)$ infer (all $i \neq j$: not(p_{i1} and p_{j1}))</p> | <p>branching conditions in any specification are deterministic</p> |

State Transformation Rules:

- | | |
|--|--|
| <p>E.1 TRUE infer M_0: elab</p> | <p>M_0 specifies a main program (usually an Ada procedure). Elab specifies elaboration as defined in the ALFM.</p> |
| <p>E.2a M: elab infer ∇C_i: elab</p> | <p>M is an Ada module, C_i is any component of M requiring elaboration.</p> |
| <p>E.2b (All C_i : C_i : elab ∇) infer ∇M: init</p> | <p>Elaboration of components leads to satisfying initial conditions.</p> |
| <p>E.2c C_i : elab infer ∇C_i .init</p> | <p>Elaboration of component leads to satisfying task local initial condition</p> |
| <p>E.2c C_i : init infer ∇C_i :property</p> | <p>Elaboration of component leads to satisfying task property</p> |
| <p>D.1 $\nabla \text{delay}(\delta)$ and $T \leq \text{STATE.time} < T + \delta$ infer $\text{delay}(\delta) \nabla$ and $\text{STATE.time} \geq T + \delta$</p> | <p>After a $\text{delay}(\delta)$, the new STATE.time is $\geq T + \delta$</p> |
| <p>R.1a $\nabla C:S.E$ infer $C:S.E$ and $S.\text{STATE.Q}(E)'\text{new} =$ $S.\text{STATE.Q}(E) + \text{callrec}(C)$</p> | <p>Being before the satisfaction of entry call $S:E$ by a client C leads to execution of the call and the enqueueing of the call record in the server S entry queue (cf. semantic action [write...] in = $\text{const}(C:S.E)$ in ch. 3.</p> |

- | | |
|---|---|
| <p>R.1b ∇ S:E and S:ECOUNT > 0 infer S:E and inx(S'STATE.Q(E).CallerId, S, E)</p> | <p>Being before the satisfaction of an accept when an entry queue is non-empty leads to a rendezvous between server S and client identified by the E queue call record.</p> |
| <p>B.2 ∇ OR(C: S:E; delay(δ), timeout and P) and C'STATE.time = T) infer inx(C,S,E) and C'STATE.time < T + δ or seq(delay(δ) ∇ , timeout and P) and C'STATE.time \geq T + δ.</p> | <p>(timed entry call) Either the client call leads to a rendezvous with the server or the time delay is completed.</p> |
| <p>B.3 ∇ OR(S:E ; delay(δ), timeout and P) and S'STATE.time = T) infer inx(S'STATE.Q(E).CallerId,S,E) and S'STATE.time < T + δ, or seq(delay(δ) ∇ , timeout and P) and S'STATE.time \geq T + δ</p> | <p>(timed rendezvous) Either a rendezvous between tasks S and C occurs or the time delay is completed.</p> |
-

9.3 Proof Structure

In this section, we give the structure for verification of system properties with the $\forall\Box$ quantifier. The proof structure consists of the following parts:

- 1) Show that the initial state
 leads to one or more characteristic states.
- 2) Show that each characteristic state
 leads to another characteristic state.
- 3) Show that the system property holds for each state
 of the finite model determined in steps 1 and 2.

In showing steps of progress of tasks, operations local to tasks can be interleaved in unknown order. Task behaviors are synchronized at points of rendezvous. We show this by grouping proof steps in the form of a concurrency map [Sto 89]. For example, with two tasks U and V, proof steps are organized in the style shown below. The proof lines

are interpreted as representative of equivalent reorderings of the lines. Two proof lines can be interchanged provided (i) the lines identify states of different tasks and (ii) the lines are not separated by a synchronization lines shown as "=====". A synchronization line indicates that states below the line cannot be interleaved with states above the line. Thus, lines 1 through 5 can be interleaved in any order which preserves the U1,U2,U3 ordering, the V1,V2,V3,V4 ordering, and the V3,U3 ordering.

| line no. | U state | V state | state pred | justification |
|----------|---------|---------|------------|---------------|
| ===== | | | | |
| 1 | U1 | | p1 | some-rule |
| 2 | | V1 | p2 | next-rule |
| 3 | U2 | | p3 | rule |
| 4 | | V2 | p4 | " |
| 5 | | V3 | p5 | " |
| ===== | | | | |
| 6 | U3 | | p6 | " |
| 7 | | V4 | p7 | " |
| ===== | | | | |

10. Example Proof of a Specification

The system property for WalkLightSystem in Figure 1b is a safety property. It requires that only walkers going in the "Green direction" can perform the Walk operation. This mutual exclusion property is achieved by the tasking system protocol, but that is not proved by the individual task specifications. In general, a WalkLightSystem can have $n \geq 2$ Walker tasks. However, for the proof example, we show only 2 Walker tasks, represented as $W(i)$ and $W(j)$. The proof extends in a straightforward way for multiple Walker tasks. In general, there are three characteristic states: (1) all W tasks in the "Green" direction, (2) all W tasks in the "Red" direction, and (3) some W tasks in the "Green" direction and some W tasks in the "Red" direction. The following proof treats only case (3).

Proposition (Walklight System System Property)

From $W(i)$, $W(j)$, and L (properties for the i th and j th walkers and LightController, respectively) infer

$$\forall \square (\forall i, j: i \neq j: \text{Walker}(i).\text{Direction} \neq \text{Walker}(j).\text{Direction}) \\ \Rightarrow \text{not } (\text{Walker}(i):\text{Walk and Walker}(j):\text{Walk}))$$

Proof (Walk Light System Specification: multiple walkers, opposite directions case).
Let W_k represents the WalkLightSystem, $W(i)$ and $W(j)$ represent tasks of WalkerType, and L represents the LightController. $\emptyset i, \emptyset j, \emptyset L$ represent the task properties for $W(i)$, $W(j)$, L . States are numbered as $n, 1:n, 2:n, 3:n$ for the n th state of $W_k, W(i), W(j), L$, respectively.

Part 0 (elaboration steps) The case where there are multiple walkers going in opposite directions, the characteristic state is when walkers $W(i), W(j)$ are either walking or calling the controller pushbutton and the light controller is before its PushButton entry. This state is characterized by the following predicate:

$$\nabla L:\text{PushButton and} \\ (\forall k: W(k):\text{Walk and Direction} \neq \text{RedDirection}) \\ \text{or} \\ W(k): L.\text{PushButton}(\text{RedDirection}))$$

In the following proof, we consider only the case where one walker walks while another walker enters into rendezvous with the LightController PushButton entry; this form of the characteristic state is expressed in terms of the following predicate:

$$\text{AND}(\nabla L:\text{PushButton and RedDirection}=\text{NS}, \\ \nabla W(i):\text{Walk and Direction} = \text{EW}, \\ \nabla W(j):L.\text{PushButton}(\text{NS})).$$

State

| | 1: | 2: | 3: | | | |
|-----|----|------|------|---|-----------|---------------|
| | Wk | W(i) | W(j) | L | Condition | Justification |
| 0.1 | | | | | Wk: elab | E.1, sys elab |

| | | | |
|-------|-----|--|--|
| 0.2 | | Wk: init▼ and (Light = (Red, Red)) | E.2b, P.1a prog. |
| 1 | 3:0 | L: elab | E.2a, comp elab |
| | 3:1 | ▼ ϕ L | E.2c, comp elab |
| | 3:2 | L: init▼ and (RedDirection = NS) | E.2, comp elab |
| 2 | 1:0 | W(i): elab | E.2a, comp elab |
| | 1:1 | ▼ ϕ 1 | E.2c, comp elab |
| 3 | 2:0 | W(j): elab | E.2a, comp elab |
| | 2:1 | ▼ ϕ 2 | E.2c, comp elab |
| ===== | | | |
| 4 | | Wk: elab ▼ | P.1, local prog. |
| | 3:3 | ▼ L: delay (MinWalkTime) | 3:1, init; assume |
| | 3:4 | delay(MinWalkTime) ▼ | D.1, local prog. |
| | 1:2 | W(i):SelectDirection(Direction)▼ and Direction = EW | P.1, local prog. assumed EW WLOG |
| | 1:3 | W(i):Walk | P.1, steps 3:2, 1:2 |
| | 3:5 | not IsClear(EW) | fr 1:3 |
| | 2:2 | W(j):SelectDirection(Direction)▼ and Direction = NS | P.1, local prog. assumed NS WLOG |
| | 2:3 | ▼ W(j):L.PushButton(NS) | fr 3:2,2:2, B.1,prog (rendezvous prog.) |
| | 1:4 | W(i):Walk ▼ | P.1b, loc. prog |

Part 1:

Discussion: A characteristic state is $\langle 4, 1:3, 2:3, 3:5 \rangle$. W(i) sees green and walks while progress by W(j) is blocked pending rendezvous with the chosen button. In this state, the system property is satisfied, since no rendezvous with the LightController has occurred (basis step).

| | | |
|-----|----------------|----------------------|
| 3:6 | IsClear(EW) | fr 1:4, assumed WLOG |
| 3:7 | L'3:6.time = T | assumed WLOG |

| | | |
|-------|--|--|
| 3:8 | PushButton(NS)'COUNT > 0 | fr 2:3 |
| 3.9 | ▼ seq(L: PushButton, L: ChangeLights, ...) | fr 3:6, 3:8, specif. |
| 3.10 | seq(▼ L: PushButton, ...) | fr 3.3, seq prog. |
| 3.11 | inx(W(j), L, PushButton) | fr 3.10 & B.3 |
| 3.12 | L'STATE.time < T + WaitTime | fr 3.11, B.3, specif. (<i>timed rendezvous</i>) |
| 3.13 | ▼ L: ChangeLights and RedDirection = EW | fr 3.11, P.2 prog. |
| 3.14 | ▼ L: FlushQueue | fr 3.13, P.2 prog. |
| 3.15 | L: FlushQueue ▼ | fr 3.14, P.1 |
| 3:16 | PushButton(NS)'COUNT = 0 | fr 3:15 |
| 3:17 | IsClear(NS) | fr 3:16, assume WLOG |
| ===== | | |
| 2:4 | ▼ W(j):Walk | fr 3:14, prog. |
| 2:5 | W(j):Walk ▼ | fr 2:4, P.1b, prog |
| 2:6 | W(j):SelectDirection(Direction)▼ and Direction = NS | Rule P.1, local prog. new green direction |
| 2:7 | | |
| 2:7 | ▼ (SeeGreen & W(j):Walk) | fr 2:6, 3:17 |
| 3:18 | ▼ L: delay (MinWalkTime) | fr 3:15, P.1, prog. |
| 3:19 | delay(MinWalkTime) ▼ | fr 3:18, D.1, delay |
| 1:5 | W(i):SelectDirection(EW)▼ and Direction = EW | Rule P.1, local prog. new red direction |
| 1:6 | ▼ W(i):L.PushButton(EW) | B.1, 3:8, det. branch |
| 3:16 | PushButton(EW)'COUNT > 0 | fr 1:6 |
| 3:17 | L'17.time = T | assumed WLOG |
| 3:18 | ▼ L: PushButton(EW) | fr 3:17, 3:16, B.3 |
| 3.19 | L'19.time < T + WaitTime | fr 3:18, B.3 (<i>timed rendezvous</i>) |
| ===== | | |

Discussion: the system state <4, 1:6, 2:7, 3:18> is a characteristic state which satisfies the following predicate:

AND(▼ L:PushButton and RedDirection=EW,
 ▼ W(j):Walk and GreenDirection = NS,
 ▼ W(i):L.PushButton(EW)).

This predicate is symmetric with the earlier predicate used to characterize system state $\langle 4, 1:3, 2:3, 3:5 \rangle$. Now walkers going NS will eventually walk while EW walkers wait. Again in this state, the system property is satisfied, since no rendezvous with the LightController has occurred. The system property is also satisfied in each of the intermediate states, since all $W(j)$ wishing to walk in the NS direction are blocked while all $W(i)$ walkers going EW progress. Thus the general computation tree folds into a transition graph, which reduces the system behavior to a finite state behavior.

Part 2:

1. For all states in the Part 0 proof, no $W(i):Walk$ operation is yet started. Hence, the system property is trivially true.
2. For all states in the Part 1 proof, only the set of EW walkers, arbitrarily selected as $W(i)$, can walk. The $W(j)$ NS walkers are blocked in the $W(j):L.PushButton$ call. Again, the system property holds. Hence, we have shown that the desired system property holds over each path from the initial state to the characteristic state, and in all states between a characteristic state and its recurrence.

11. Experimentation Layer

During the experimentation phase, prototypers use the workbench to perform testing of a prototype which have been synthesized from either the visual or specifier's layer. The Experimentation Layer provides simulators which offer a realistic testbed for prototype systems. A simulator is a model of the environment in which an RTS is expected to live. The workbench can also be used to construct these models of the environment. The simulators are used in the experimentation layer to exercise the system modules as well as test error recovery paths. The form of rapid prototyping provided by this workbench encourages "software" reusability. Rapid prototyping of system software utilizes the AdaSage approach (i.e., libraries of existing automata and temporal specifications can be used as building blocks in constructing a new, real-time system). In addition to reusing existing program code, the workbench encourages the

reuse of concepts represented by archived automata and temporal formulas. This is in keeping with the form of reusability found in [Chi90].

12. Conclusion

The purpose of the proof system is to provide the basis for developing proofs of consistency for Ada / TL_{rt} tasking system specifications. These proofs would be written with the help of the Specifier's Assistant of the RTS Prototyper's Workbench. As a result of proving the consistency of a system property relative to local task properties, the code for task bodies would be extracted as a byproduct of the constructive proof of a specification. The proof technique reduces an infinite computation tree exhibiting system behavior to a transition graph with a finite number of states. The key to this reduction is the identification of certain system states which we call characteristic states, i. e. recurring states. Finally, the proof system for Ada/TL_{rt} introduces the algebra, axioms and rules useful in specifying tasking system behavior. Preliminary work has been done on the first two layers of the Prototyper's workbench [Pet91a, Pet91b, Pet91c, Pet90a, Pet90b, Ram 91]. The development of the experimentation layer is part of the future work we envision for the Prototyper's Workbench.

References

- [Alp 86] Alpern, B.L. Proving Temporal Properties of Concurrent Programs: A Non-Temporal Approach, Ph.D. dissertation, Cornell University, TR-86-732, 1986.
- [Alu90] Alur, R., and D. Dill (1990). Automata for modeling real-time systems. *Automata, Languages and Programming*. Lecture Notes in Computer Science 443, Springer Verlag, NY. Pp. 322-335.
- [Bar82] Barringer, H. Mearns, I. Axioms and Proof Rules for Ada Tasks. *IEE Proceedings, Part E: Computers and Digital Techniques*, vol. 129, No. 2 (Mar 1982), 39-48.
- [Bar86] Barringer, H. Mearns, I. A Proof System for Ada Tasks. *The Computer J.*, vol. 29, No. 5 (Oct. 1986), 404-415.
- [Ben81] Ben-Ari, M. Manna, Z. Pnuell, A.. The Temporal Logic of Branching Time. *Proceedings of the 8th ACM Symposium on Principles of Programming Languages*. Jan. 26-28, 1981, 164-175.
- [Bur87] Burns, A. Lister A. Wellings, A. *A Review of Ada Tasking*, LNCS 262, Springer-Verlag, 1987.
- [Chi90] Chin, S.-K. Formal methods and Reusable Software. *Proceedings of the 5th Annual Knowledge-Based Software Assistant (KBSA) conference* (Sept., 1990), 42-56.

- [Cla85] Clarke, E. Browne, M. Emerson, E. Sistla, A. Using Temporal Logic for Automating Verification of Finite State Systems. In *Logics and Models of Computer Science*. NY: Springer-Verlag, 1985, 3-26.
- [Con 86] Constable, R.L. et al. *Implementing Mathematics with the Nuprl Proof Development System*. Prentice-Hall, NJ, 1986.
- [Cro88] Cronhjort, B. (1988). Specification and Quality Assurance of Real-Time Software. *IFAC/IFIP Symposium on Software for Computer Control*, pp. 9-14.
- [DoD83] U.S. Dept. of Defense, *Reference Manual for Ada Programming Language*, ANSI/MIL STD 1815A-1983 NY: Springer-Verlag, 1983.
- [Hand88] Handelsmann, D.A. and Stengel, R.F. (1988). Perspectives on the Use of Rule-Based Control. *Proceedings of the IFAC Workshop on Artificial Intelligence in Real-Time Control*, pp. 27-32.
- [Han90a] Hankley, W., and J.F. Peters (1990a). Temporal Specification of Ada Tasking. *Proceedings of the 23rd Hawaii International Conference on System Sciences: Software Track*, Vol. II, pp. 410-419.
- [Han90b] Hankley, W., and J.F. Peters (1990b). A Proof Method for Ada/TL. *Proceedings of the 8th ACM-IEEE Annual National Conference on Ada Technology*, pp. 392-399.
- [Har 90] Harel, E. et al. Explicit Clock Temporal Logic, *Proceedings of the 5th Annual IEEE Symposium on Logic in Computer Science*, , 402-413, June, 1990.
- [Hen 91] Henzinger, T.A., et al. Temporal Proof Methodologies for Real-Time Systems, *Proceedings of the 18th Annual ACM Symposium on Principles of Programming Languages*, 353-366, Jan., 1991.
- [Jac85] Jackson, M. Developing Ada programs using the Vienna Development Method. *Software--Practice and Experience*, vol. 15, no. 3, March 1985, pp. 305-318.
- [Jon90] Jones, C.B. *Systematic Software Development Using VDM*. NJ: Prentice-Hall, 1990.
- [Kop88] Kopetz, H., and K. Kim (1988). Consistency constraints in distributed real time systems. *Proceedings of the 8th IFAC Workshop on Distributed Computer Control Systems*, pp. 29-34.
- [Kop89a] Kopetz, H. (1989a). Real-Time Systems. Research Report Nr. 12/89, Institut für Technische Informatik, Technische Universität, Wien, Austria.
- [Kop89b] Kopetz, H. (1989b). Design of Real-Time Computing System. *Proceedings of Joint Univ. of Newcastle Upon Tyne/International Computers Limited Seminar*, pp. IV.25-IV.58.
- [Krö 85] Kröger, F. Temporal Logic of Programs Lecture Notes, Report TUM-18521, Institut für Informatik, Technische Universität München, 1985.
- [Lam83] L. Lamport, "Specifying Concurrent Program Modules", *ACM Trans. on Programming Languages and Systems*, vol. 5, no. 2, April 1983, 190-222.
- [Man 81] Manna, Z. Pnueli, A. Verification of Concurrent Programs, Part 1: the temporal framework, Report No. STAN-CS-81-836, Dept. of Computer Science, Stanford University, June, 1981.
- [Man 83a] Manna, Z. Pnueli, A. Verification of concurrent programs: a temporal proof system, Report No. STAN-CS-83-967, Dept. of Computer Science, Stanford University, June, 1983.
- [Man83b] Manna, Z. Pnueli, A. How to Cook a temporal proof system for your pet language. *Proc. ACM POPL*, Austin, TX, (Jan. 1983), 141-154.
- [Mil 89] Milner, R. *Communication and Concurrency*. Prentice-Hall, Inc., NJ, 1989.

- [Ost 89] Ostroff, J.S. *Temporal Logic for Real-Time Systems*. John Wiley & Sons, inc., New York, 1989.
- [Ost 90] Ostroff, J.S. Wonham, W.M. A Framework for Real-Time Discrete Event Control, IEEE Transactions on Automatic Control, vol. 35, no. 4, 386-396, Ap., 1990.
- [Pet 90a] Peters, J.F. Constructive Specification of Communicating Processes Using Temporal Logic, Ph.D. dissertation, Computing & Information Sciences, Kansas State University, 1990a.
- [Pet 90b] Peters, J.F. and Hankley, W. Proving Specifications of Tasking Systems Using Ada/TL, Proceedings of ACM Tri-Ada'90, 4-13, Dec., 1990b.
- [Pet 91a] Peters, J.F. and Ramanna, S. (1991a). Modelling Timed-Behavior of Real-Time Systems with Temporal Logic. To appear in Cybernetics and Systems: An International Journal, 1991.
- [Pet 91b] Peters, J.F. and Ramanna, S. (1991b). Constructing Real-Time Systems from Temporal I/O Automata. Report No. SU-CIS-91-22, School of Computer & Information Science, Syracuse University.
- [Pet 91c] Peters, J.F. and Ramanna, S. (1991c). Prototyping Provably Correct Real-Time Systems, Report No. SU-CIS-91-23, School of Computer & Information Science, Syracuse University, July, 1991.
- [Pnu79] Pnueli, A. The Temporal Semantics of Concurrent Programs. In *Lecture Notes in Computer Science 70*, Springer-Verlag, NY, 1979, pp. 1-20.
- [Pnu82] Pnueli, A. DeRoeve, W. P. Rendezvous with Ada--A Proof Theoretical View. *Proc. AdaTEC Conference on Ada*, Arlington, VA, Oct. 6-8, 1982, 129-137.
- [Ram91] Ramanna, S. and Peters, J.F. Explicit Clock Temporal Logic in Constraint Checking for Real-Time Systems. To appear in the Proceedings of IFAC AIRTC'91, September, 1991.
- [Sto89] Stone, J. A Graphical Representation of Concurrent Processes. *Sigplan Notices*, vol 24, no 1 (Jan 1989), pp 226-235.
- [Wen90] Wengelin, D., and L. Asplund. Application of Ada on a Distributed Missile Control System. Proceedings of ACM Tri-Ada'90 (Dec., 1990), 300-312.
- [You85] Young, W.D. et al. Proving a Computer System Secure. *Scientific Honeyweller*, vol. 6, no. 2 (July 1985), 18-27.

EXPERIMENTAL EVALUATION OF OPTICAL SWITCHING TECHNOLOGIES

Dr. Dean Richardson

Abstract: During my 10-week tenure at the USAF Photonics Center, I was involved with two distinct though related projects. First, I helped align and optimize a mode-locked picosecond dye-laser system for use in studying cascability of Self-Electro-Optic-Effect Devices (SEEDs) on loan from Bell Laboratories. Second, I participated in experimental design, optical alignment, and data acquisition efforts aimed at characterizing the reflection-mode behavior of a novel class of tungsten-oxide thin-film optical switches. Since my previous "hands-on" experimental experience had been somewhat limited, both of these assignments enhanced my ability to perform confidently and capably in a laboratory environment.

Introduction:

In recent years, a variety of schemes for demonstrating "optical switching" have been proposed, including semiconductor etalons, SEEDs, waveguides, etc.. One of the objectives of the US Air Force Rome Laboratories Photonics Center is to understand and evaluate the performance and utility of potential optical switching devices, with an eye to their possible application in both digital and analog optical computing architectures. As part of this ongoing effort, the Optical Switch Evaluation Facility (OSEF) has been

set up. The facility consists of an impressive array of continuous-wave and short-pulse laser systems as well as associated optical components, detectors, signal processing electronics, micropositioning hardware, and related computer-interfacing equipment/software.

My assignment as an RDL summer visiting scientist was to work with Mr. Joe Osman, who has primary responsibility for the day-to-day operation of the OSEF. After discussing my interests and objectives for the summer upon my arrival, we decided that three OSEF projects would be particularly appropriate for me to be involved in: 1) etalon bistability studies in collaboration with researchers at Georgia Tech (my Ph.D. work had centered on etalon-based switching, including bistability); 2) SEED analysis and characterization (SEEDs were available for testing through a cooperative agreement with Bell Labs; I had researched SEED performance characteristics and fundamental operating principles as part of my dissertation); and 3) measurements of intensity-dependent refractive-index behavior in nonstoichiometric tungsten-oxide photorefractive thin films, in collaboration with Prof. Joseph Chaiken of Syracuse University.

Each of the three projects required the use of a different laser system or subsystem available in the OSEF. The etalon studies required the use of the Argon-ion-pumped titanium-sapphire cw laser (tunable in the vicinity of the GaAs bandedge) since cw bistability measurements were contemplated in GaAs/AlGaAs etalons. The SEED experiments were to involve picosecond-pulsed excitation of SEED elements and arrays, again in the 840-nm region of the spectrum; thus a dye laser pumped by a second (mode-locked) Argon-ion laser was to be used. The tungsten-oxide switch characterization did

not initially require ultrashort pulses; as a result, the samples were to be studied using 100-ps FWHM pulses from the mode-locked Ar:Ion laser operating at 514 nm.

As I arrived at the Photonics Center for the summer, only the two Ar:Ion pump lasers were operational; both the dye and Ti:Sapphire lasers needed to be realigned and integrated within their respective optical trains for device testing. In addition, the appropriate sample rotation set-up for reflection-mode studies of the tungsten-oxide devices was not yet complete. With my experience rooted firmly in theoretical simulation, it was obvious that I would be "getting my hands dirty" in the lab. Indeed, it quickly became apparent that the summer would provide me with a sort of "crash-course" in both the frustrations and rewards of laser-based optical experimentation.

Statement of Problem(s):

As alluded to above, three basic tasks faced us at the outset of my tenure at the Photonics Center:

- 1) Align and optimize the Ti:Sapphire laser to produce tunable cw output in the vicinity of the GaAs bandedge. This would involve dealing with cooling water flow problems to the Ar:Ion pump laser as well. Search for excitonic optical bistability in samples grown by researchers at Georgia Tech.
- 2) Align and optimize the dye/Ar:Ion laser system to produce picosecond pulses; design an optical train to analyze the SEEDs; test SEED input/output, switching, and cascading behavior.

- 3) Adapt and refine an existing single-wavelength pump-probe setup to allow reflection-mode characterization of tungsten-oxide thin-film switches using 100-ps mode-locked pulses. From graphs of reflected intensity as a function of angle at different intensities, evaluate the photorefractive behavior of the films.

Results/Narrative of Work Performed:

To be blunt, little if any publishable data was gathered as a direct result of the work I performed during the summer; therefore I have chosen to summarize my work in a narrative fashion, describing the problems we had, the ways we dealt with them, and our eventual successes.

Etalon studies:

Before Ti:Sapphire alignment could proceed, cooling water pressure/temperature fluctuation problems with the (initially) Spectra-Physics Ar:ion pump laser had to be resolved. Joe Osman and Tech. Sgt. Brian Longacre had envisioned some sort of reservoir system coupled with a heat exchanger; this would provide a stable source of water at the required temperature. Not long after early efforts to peak the Ar:ion power had begun, the tube cracked, necessitating the swapping of the Spectra-Physics laser for a Coherent laser which was then not in use. This laser suffered from the same cooling water problems the Spectra-Physics Ar:ion had had, so the search for a reservoir/heat-exchanger intensified.

Joe, Brian, and I went to talk to an experienced problem-solver in the high-power electronics lab at Griffiss AFB; he had extensive expertise with cooling and pumping systems. A series of discussions led to several ideas for plausible solutions to our problem; actual hardware was a bit harder to come by. We soon turned (partly in frustration, partly because of inevitable lead times in procuring needed components) to the next project. At this point my involvement in the etalon work stopped. This work was taken up again later in the summer by TSgt. Longacre, who eventually got the enhanced cooling system in place. As of this writing, the actual etalon testing has not yet begun, pushed aside by the urgent desires of other groups to use the Ti:Sapphire laser for their purposes.

SEED testing:

While the etalon effort was on hold, Joe, Brian and I turned our attention to aligning and optimizing the picosecond (ps) dye laser system in preparation for SEED performance measurements. Conversations with Bob Morgan of AT&T and Prof. Joe Chaiken of Syracuse University had already produced a working list of SEED-related experimental objectives; our task now was to achieve mode-locked, synchronous-pumped generation of good-quality ps pulses.

1) Initial Goals:

From an alignment standpoint, the OSEF's dye-laser system seemed to me to have three basic components: the Ar:Ion pump laser, the dye laser, and its associated

extended cavity (needed for sync-pumping purposes). The pump needed to be successfully mode-locked so as to produce pulses of approximately 100 ps full-width at half-maximum (FWHM). These pulses then were to be directed into the dye-stream, which in turn would provide a gain medium capable of generating pulses of only several ps FWHM. The image of the pump/dye-stream fluorescence was then to be focused and optimized so as to produce lasing within an extended cavity, whose length would be matched to that of the pump laser, allowing "sync-pumped" operation. Each of the three components presented their own unique challenges and pitfalls.

2) CW alignment

Brian and I began by substituting a recently-purchased sapphire dye-nozzle for the less precisely-machined nozzle originally supplied with the dye-laser, with the goal of achieving greater dye-stream stability, as well as avoiding the "moustaching" phenomenon common to dye-jets. As a zeroth-order test of the system alignment, we attempted to observe lasing of the extended cavity while pumping the dye-laser using the cw "all-lines" output of the Argon-Ion laser. Initially, I watched over Brian's and Joe's shoulders as they tweaked the various turning- and intracavity mirrors used to guide the pump beam into the dye laser, focus it on the dye jet, and image the induced fluorescence. Before long, they had me making the mirror adjustments myself. By relying on a surprisingly low-tech measurement device, namely one of Joe's business cards, we soon had the various reflected and transmitted spots aligned. After a tedious effort to align the extended cavity output mirror (for which transverse positioning was

accomplished by hand, instead of with thumbscrews) we were able to observe cw laser output from the dye laser at 840 nm.

The output was not as strong as had previously been observed in cw operation using this same set-up; Joe concluded that the Styryl-9 dye being used as a gain medium had aged sufficiently so as to merit replacement. Brian and I disposed of the old dye and mixed up a new batch in hopes of "recharging" the gain medium's performance. With the new dye pumped into the system, we noticed a modest though definite increase in dye-laser output; the improvement was significant enough to allow us to proceed to the next step, namely mode-locking the pump laser.

3) First try at mode-locking:

Superficially, the switch from cw to mode-locked operation was basically a matter of swapping the "all-lines" mirror for a corresponding mode-locking head. Of course, the actual transfer of energy from the various undesired modes to the single desired one required the use of a servo-mechanism and associated electronics; as I understand it, these serve to periodically modulate the pump-laser cavity length by adjusting Bragg-resonances in a specially-designed crystal. In practice, I simply adjusted the driving frequency on an electronic stabilization module while simultaneously observing the "locking" process on a meter (which tracked the servo/crystal interaction) and on a Ghz oscilloscope (which when connected to a fast-risetime detector, displayed the gradual build-up of mode-locked pulses, and their temporal behavior).

Once I was sufficiently comfortable with this process so as to be able to get the

pump laser up and running mode-locked each morning, we turned again to the dye-laser and its extended cavity. Mode-locking significantly reduced the average output power coming from the pump laser and changed its alignment, so the two turning mirrors and the various dye-laser mirrors needed to be re-adjusted. Having again optimized the overlap of the various reflected and transmitted spots, we sought once more to manually position the extended cavity output coupler to allow lasing. Following a great deal of repeated trials, the extended cavity finally lased, producing (after extensive and repeated optimization of each mirror in turn) an average of 400 mW of output power at 840 nm in mode-locked operation.

4) Second try at mode-locked sync-pumped operation:

We had hoped to obtain at least 1 watt of sync-pumped output to allow for subsequent losses in the optical train to be used for testing SEEDs. After much thought concerning possible means of further enhancing the extended-cavity alignment, it was decided that a more careful optimization of both the focus and shape of the dye-jet fluorescence images was needed. Additionally, I suggested putting the extended cavity output coupler on an x-y translation stage with micrometer-type adjustment screws, to allow smooth, repeatable positioning of the end-mirror.

Though the translation mount gave greater flexibility and accuracy in positioning the end-mirror, it also increased the mirror's height relative to the dye laser, requiring us to raise the dye-laser and readjust the input pump beam as well. After the dye-laser and pump beam had been carefully leveled, the spot alignment procedure was repeated.

Unfortunately, with the mid-cavity fluorescence image optimized for maximum power, we found that the transmitted spot at the end-mirror plate missed the extended-cavity output coupler by a wide margin. Even with the translation stage adjusted to the maximum of its possible travel, only a small portion of the spot struck the output coupler. Further, since our optical table had tapped holes at 1" spacings, while the translation stage provided a maximum travel in either direction of only 0.25", we could not resolve the spot misalignment by simply moving the translation stage over to the next row of holes. We tried translating the dye laser itself, but this introduced the same non-repeatability we had earlier experienced with the end-mirror, since the dye-laser positioning feet were no longer locked into the 1"-spaced tapped table holes. The slightest tap on the dye-laser would destroy the optimum alignment. Additionally, there remained the problem of centering the pump beam on the second turning mirror, which also suffered from the 1"-step movement problem.

5) An additional degree of freedom

Repeated attempts failed to find a suitable arrangement of the turning mirror, dye laser, and end mirror which would permit lasing yet allow the various components to be securely locked into table holes instead of simply being clamped down after tedious hand-adjustment. Finally we decided to introduce an additional degree of freedom by putting the second turning mirror on a two-dimensional translation stage. Now the dye laser feet could be locked into the appropriate rows of table holes, the end-mirror could be adjusted to the center of its transverse travel, and the pump-beam mirror could be adjusted

accordingly to allow the fluorescence image to be optimally focused on the output coupler. After meticulous tuning of the vertical, horizontal, and pivot screws of each intracavity mirror, we again achieved lasing of the sync-pumped cavity, this time with approximately 1.2 W of output power.

This represented a substantial improvement over previous mode-locked results with this system; feeling that we were ready to proceed to the next step, we prepared the autocorrelator assembly to accept the pulsed dye-laser output, hoping to find that we indeed had pulses as short as a few ps FWHM. After attempting to insert an enclosure tube between the dye laser and the end-mirror to mechanically stabilize the extended cavity, we were unable to easily regain lasing behavior. Apparently the dye laser housing had been jarred slightly during the tube-insertion process. Puzzled that the dye-laser should become so seriously misaligned so easily, we made very fine adjustments to the various mirror adjustment screws so as to determine which one was particularly sensitive.

The main mirror adjustment screws in the dye-laser tighten their respective mirror mounts down against one of several quartz rods which run the length of the dye-laser and provide mechanical stability for the longitudinal cavity. In earlier tightening, one of the rods had become cracked. This had not seemed to cause any major difficulty in output power optimization, so we had not been overly concerned at the time. It was now apparent that with the rod broken into several pieces, the mirror orientations referenced to it had become highly unstable. Though we could probably come close to our best observed output power by careful re-optimization, we would have no assurance that we

could maintain that alignment long enough to perform any SEED characterization. We decided to order a replacement quartz rod assembly and pursue other projects until the rod arrived. By the time it came about three weeks later, we had become heavily involved in optical system design for studies of photorefractive tungsten-oxide switches, to be performed in collaboration with Prof. Joe Chaiken of Syracuse University.

Characterization of photorefractive thin-film nonlinear interface switches:

Soon after our realization that the dye-laser quartz rod needed replacing, Prof. Chaiken brought us samples from his lab in Syracuse of non-stoichiometric tungsten-oxide thin films, grown in a "bullseye" pattern on small quartz substrates. Prof. Chaiken felt confident that the films would exhibit nonlinear (i.e. intensity-dependent) refractive index changes at sufficiently high input powers; our job was to refine the existing optical train to allow detailed studies of reflected intensity vs. angle. By comparing the angular dependence of the reflected signal from the thin-film sample with that from an uncoated quartz substrate -- particularly near Brewster's angle -- we would be able to determine the sign and magnitude of the optically-induced refractive index changes that might be a sign of nonlinear switching.

An optical train for use in sample characterization had been set up before my arrival. It included dual optical delay lines and a Klinger x-y-z positioning set-up with a goniometer attached. Each of these could be manually translated/rotated using front-panel switches on the Klinger controller banks; through an IEEE interface, they could also be controlled by computer. Brian DeVaul, a student working for Joe for the

summer, had been assigned to develop several menu-driven programs to help control the positioning equipment as well as other devices. Until they were completed, I would operate the devices from the control panel.

A great deal of system design and alignment needed to be completed before actual data-taking could begin. Several weeks were spent centering the sample over the rotation axis of the goniometer, aligning both the pump and probe beams (generated from the Argon-Ion output using a beamsplitter) on the sample, and adjusting the polarization appropriately so as to make Brewster's angle measurements possible. During this time, my main task was to control the sample rotation so as to verify that the input spots stayed centered on the sample face. I also assisted in choosing the polarization scheme we finally employed.

After these initial alignment chores had been completed, Joe Osman constructed an impressive boom arrangement so a camera or detector head could rotate on the same axis as the sample, at twice its angular rotation rate. This allowed us to measure the reflected intensity as a function of angle, in a continuous manner, under computer control. As Brian DeVaul's device driver programs became operational, my role shifted somewhat; I focussed my efforts on optimizing and maintaining mode-locked operation of the Argon-Ion pump laser, which had begun to exhibit somewhat unstable behavior.

As the summer drew to a close, data-taking efforts finally begun. We were able to obtain numerous plots of reflected intensity versus angle for both the tungsten-oxide films and the uncoated substrates. On one particular day when neither Joe Osman or Joe Chaiken could be present in the lab, Brian DeVaul and I took an extensive set of data,

varying the input intensity over a wide range in search of nonlinear reflected intensity behavior. Initial results did not seem to indicate a particularly strong nonlinearity in the refractive index at the intensities we had used. The data was still being interpreted when my 10-week stay at the Photonics Center ended. Joe Osman and Joe Chaiken are now preparing to adjust the focusing optics, in hopes of raising the peak intensity incident on the thin-film samples.

Conclusions:

My stay at the USAF photonics center provided me with a significant amount of hands-on experience with pulsed and cw lasers, mode-locking, and optical system design. I feel that this work will complement my previous theoretical background and allow me to be a more effective researcher in the future.

CLOCK SYNCHRONIZATION TECHNIQUES: A SURVEY

Final Report

AFOSR Summer Faculty Research Program

9 August 1991

Waleed Smari, Graduate Student

Department of Electrical & Computer Engineering

Syracuse University

121 Link Hall

Syracuse NY 13244 -1240

ABSTRACT:

In this report we give a summary of the research that has been done on the problem of clock synchronization in distributed computing systems. After presenting the theoretical background, we introduce a brief description of the works which attempted to solve this problem. The report is a summary of a detailed technical report submitted to Rome Laboratories surveying over one hundred papers on clock synchronization techniques to date. The reader is referred to the technical report for further details.

CLOCK SYNCHRONIZATION TECHNIQUES: A SURVEY

INTRODUCTION:

A distributed system consists of a collection of distinct processes (processors) which may be spatially separated, and which communicate with one another by exchanging messages. A system is distributed if the message transmission delay is not negligible compared to the time between events in a single process. Thus, a network of interconnected computers is a distributed system. A single computer can also be viewed as a distributed system in which the central control unit, the memory units, and the I/O channels are separate processes.

Even without errors or node failures, it will be impossible for each entity in a cooperating set to maintain the same view of their total state as the others, due to the arbitrary delays for messages updating or reporting states. Therefore, it is necessary to provide the cooperating entities with a common reference of specified accuracy. This procedure is called **synchronization**.

In a distributed system there are two principal reasons for providing synchronization:

- a) *Concurrency*: individual processes, even those that belong to independent applications, may need to enter a state of concurrency either to share access to system resources or to control access to shared data.
- b) *Cooperation*: processes that belong to the same application can cooperate for proper execution of the application.

Concurrency and cooperation between several processes require the exchange of data. Therefore, it is necessary to ensure the order of events of the various processes involved. This total ordering of events must encompass two notions of precedence: 1) Precedence between any two events in the same process; and 2) precedence of the message-sending event of a process with respect to the message-receiving event of another process.

At each site, synchronization is realized by assuring the construction of a definite order. This order may be different from the absolute temporal order (chronological) that an external observer would oversee on the set of the sites. This order complies with the precedence rules.

There are several techniques available for ordering of events. Logical clocks is one such technique as will be addressed shortly.

We model a distributed system as consisting of a set of processes that communicate by sending messages to each other. Each process has a physical (hardware) clock that is not under its control. Physical clocks are subject to a small rate of drift. A process's local time is obtained by adding the value of the physical clock to the value of a local "correction" variable.

The communication network may be assumed to be fully or partially connected. In the former case, every process can send a message directly to every other process. A process broadcasts a message to all the processes at the same time. All messages are delivered within a fixed amount of

time plus or minus some uncertainty (message delays). Messages may (or may not) be required to have unforgeable signatures (authenticated).

A time service algorithm is an algorithm that keeps a collection of clocks locally monotonic, synchronized, and adequately accurate with respect to some time standard. The relative importance of these properties depends on the purpose the time service is used. For example, accuracy is not important if the time service is only used to order events occurring within the distributed system.

A clock synchronization algorithm is used in a distributed system to synchronize clocks that can drift apart, to ensure that the maximum deviation between the clocks is bounded. A fault-free hardware clock (HC), even if initially synchronized with a standard time reference, tends to drift away from the standard over a period of time.

The fractional error in the measurement of a time interval (t_1, t_2) is bounded by a constant, ρ , called the *drift rate* of the clock:

$$\rho > \left| \frac{[H(t_2) - H(t_1)] - [t_2 - t_1]}{[t_2 - t_1]} \right| \Rightarrow \rho > \left| \frac{[H(t_2) - H(t_1)]}{[t_2 - t_1]} - 1 \right|$$

which can be written as $(1-\rho)(t_2-t_1) \leq [H(t_2) - H(t_1)] \leq (1+\rho)(t_2-t_1)$. If this condition holds, we say that the (hardware) clocks are ρ - **bounded**.

The problem of clock synchronization may involve one or more of the following:

- 1) maintaining synchronization of local times once it has been established;
- 2) establishing synchronization initially;
- 3) bringing into synchronization clocks that begin with arbitrary values (after being repaired) by reintegration.

It is worth noting here that when we say "the clock of process P" we usually mean "the clock of the processor on which P runs" since there is a one-to-one correspondence between processes and processors. Therefore, we do not distinguish between processes and processors.

In general, it is assumed that when messages are sent through a communication link, they are neither altered nor lost. Also, for any two sites P_i and P_j , the order in which P_i sends messages to P_j is identical to the order in which P_j receives them from P_i .

Let us briefly review the issues introduced above, namely time, clocks, events, ordering of events, synchronization, and so on.

BACKGROUND:

1. Time in a distributed system:

A global time base is an important requirement in a distributed computing system. It can be used to simplify the solutions to several design problems such as interprocessor communications, resource allocation, transaction processing, and checkpointing, especially in the presence of faults. It can also be used to implement features such as deadlines and timeouts that are essential for correct

operation of distributed real-time systems. In order to measure the time of occurrence of events and the duration of intervals between events, which are observed from different nodes, it is necessary to provide to the application a common time reference of specified accuracy. A global time base can be established by synchronizing the local clocks on different nodes in the distributed system.

Kopetz [86] identified a number of different standards of time that we may have to deal with. Relevant to our work are:

- *External Physical Time*: a standard of time which is based on some periodic physical (atomic) oscillation. It is a chronoscopic time without any irregularities. This time may be taken as an external reference for distributed real time systems.
- *Local Physical Time*: which is the time of the local physical clock in a node of a distributed computing system. The physical clock is a device which is based on some periodic physical oscillation (e.g. a quartz crystal) in order to measure the progression of physical time.
- *Approximate Global Time*: which is an approximation of a common time base that should be made available at each node of a distributed computing system. It is established by internal synchronization via the exchange of messages.
- *Internal Physical Time*: which is a time reference within each node of a distributed computing system. It approximates the external physical time. It can be thought of as an internal image of the external physical time at each node of the system.

The reader is referred to the section on synchronization types for further details.

2. Clocks:

We define a clock to be a mapping C from real time to clock time, where $C(t)=T$ means that at real time t the clock reads T . This is appropriate when clocks are used to measure the time at which some event occurred. (Ex: a car crossed the finish line).

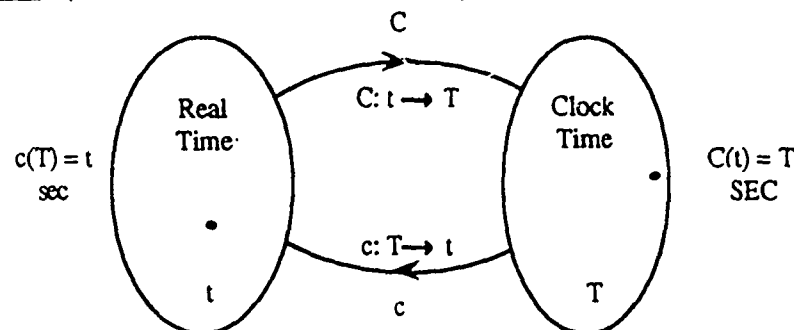


Fig. 1. The generation of clocks

Clocks are also used to determine when events are generated (Ex: a valve should be shut). In this case, it is more appropriate to define a clock to be the inverse of the function above. So it is a mapping c from clock time to real time, with $c(T)$ being the real time at which the clock has the value T .

Two clocks c and c' are said to be synchronized to within δ at a clock time T if $|c(T)-c'(T)| <$

δ . Thus, they reach the value T within δ seconds of one another. If two processes' clocks are synchronized to within δ at time T , then actions generated by the two processes at that time occur within δ seconds of one another.

Nonfaulty clocks are assumed to be monotonic. Thus, clocks (or the inverses) are represented by single-valued functions, i.e., for any real time t there is a unique T such that $c^{-1}(t)=T$.

A perfect clock is one in which $C(t) = t$ for an appropriate definition of t . A clock is correct at time t if its value $C(t) = t \pm \epsilon(t)$, $\epsilon(t)$ is the error. A clock is accurate at time t if its first derivative is one sec per sec. A clock is stable at time t if its second derivative is zero. Two clocks are consistent at time t if $|C_i(t) - C_j(t)| \leq \epsilon_i(t) + \epsilon_j(t)$, $\epsilon_i = \text{error}$. Consistency is a weaker condition than correctness. Since there is no perfect clock in the system, an algorithm can not check to see if the clocks are remaining correct, it can check to see if the clocks are remaining consistent. A clock c is a good clock during the real-time interval $[t_1, t_2]$ if it is a monotonic, differentiable function on $[T_1, T_2]$, where $c(T_i)=t_i$, $i=1,2$, and for all T in $[T_1, T_2]$, with ρ being the rate at which nonfaulty clocks can drift apart:

$$\left| \frac{dc(T)}{dT} - 1 \right| < \rho$$

3. Events Ordering and Clocks :

In a distributed system, it is sometimes impossible to say that one of two events occurred first. Therefore the relation "happened before" is only a partial ordering of the events in the system. It is assumed that a distributed system is composed of a collection of processes. Each process consists of a sequence of events. The set of events are with a priori total ordering. If the system is to meet a specification correctly, then that specification must be given in terms of events observable within the system. If the specification is in terms of physical time, then the system must contain physical clocks. Even if it does contain physical clocks, there is still the problem that such clocks are not perfectly accurate and do not keep precise physical time. Therefore, it is useful to define the "happened before" relation without using physical clocks. Logical clocks are used to do so.

To guarantee a correct system of logical clocks, we state the **clock condition** as follows: *if an event (a) occurs before another event (b), then (a) should happen at an earlier time than (b), i.e., for any events a, b: if $a \rightarrow b$ then $C(a) < C(b)$. The converse need not be true.* If the clock condition is satisfied, then: 1) For events a, b in process P_i : if $a \rightarrow b$ then $C_i(a) < C_i(b)$. 2) For events $a \in P_i$ and $b \in P_j$: if a is the sending of a message by process P_i and b is the receipt of that message by process P_j , then $C_i(a) < C_j(b)$.

Thus, to guarantee a *correct* system of logical clocks, both conditions 1 and 2 must be satisfied. Notice that (due to condition 1) when a clock is reset, it is always set forward and never back.

In order to formalize the synchronization conditions that must always be satisfied in a clock system, we introduce the following. In a distributed system, one can think of **three kinds of clocks** associated with each process P_i : real-time clock, $RT=t$, physical (or hardware) clock,

$H_i(t)$, and logical (or software) clock, $C_i(t)$.

Physical clocks can be assumed to run continuously (Lamport[85]) or have discrete "ticks" (Kopetz [86]). If the former assumption is used, then $H_i(t)$ is a continuous, monotonic, differentiable function of t except for isolated jump discontinuities where the clock is reset. The rate at which the clock is running at time t is given by $\frac{dH_i(t)}{dt}$. In order for the clock H_i to be a correct physical clock, it must run at approximately the correct rate for all t , i.e., $\frac{dH_i(t)}{dt} \approx 1$. More precisely, the drift rate of the physical clock from real time is bounded by a known constant $\rho > 0$ such that for all t : $\left| \frac{dH_i(t)}{dt} - 1 \right| < \rho \Rightarrow \frac{1}{(1+\rho)} \leq \frac{dH_i(t)}{dt} \leq (1+\rho)$.

That is, if $H_i(t)$ is the reading of the physical clock of process P_i at (real) time t , then for all $t_2 \geq t_1$: $\frac{(t_2 - t_1)}{(1+\rho)} \leq H_i(t_2) - H_i(t_1) \leq (1+\rho)(t_2 - t_1)$

The clock above is then said to be ρ -**bounded**. Typical crystal controlled clocks have $\rho \leq 10^{-5}$.

In addition, correct physical clocks must be synchronized so that, for all P_i, P_j , and t :

$$|H_i(t) - H_j(t)| < \beta, \text{ where } \beta \text{ is a sufficiently small constant.}$$

It is assumed that the N hardware clocks of processors have a much higher resolution than the time intervals which must be measured (the process-to-process communication delays). Thus, if the delays observable are of the order of msec., we assume the hardware clocks have a μ sec resolution.

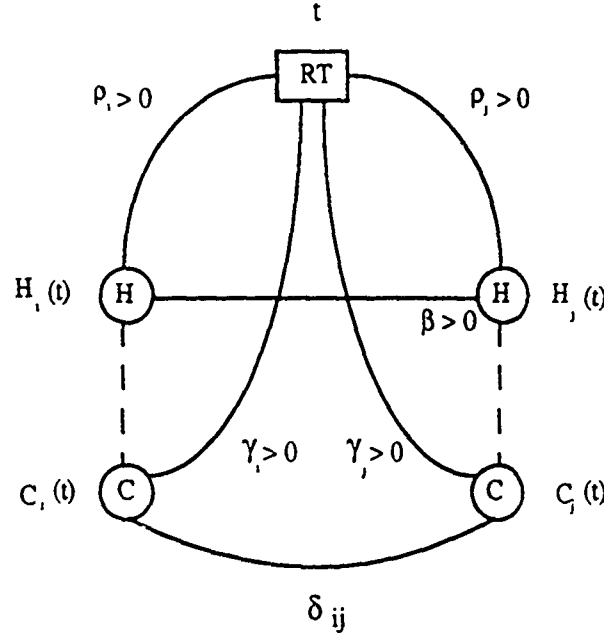


Fig. 2. Clocks and Drifts

There has been a great deal of work done on the problem of synchronizing physical clocks, as we shall see later.

A software synchronization protocol maintains approximately synchronized *logical clocks* in a

distributed system. The protocol maintains on each correctly functioning processor P_i a logical clock C which measures the passage of real time with an accuracy comparable to that of a hardware clock and which, at any time t , displays a clock time $C_i(t)$ that is within some known bound δ_{\max} from the clock times displayed by all other logical clocks $C_j(t)$ running on another correct processor P_j :

$$\forall t: |C_i(t) - C_j(t)| < \delta_{\max}$$

A logical clock is maintained by periodically computing an adjustment function for a hardware clock as shown in figure 2.

CLOCK SYNCHRONIZATION:

In a distributed computing system it is often required to measure the duration between two events which have been observed by two different nodes or to identify and record the time of occurrence of a distributed event, i.e., an event which can be observed from different nodes in a distributed system. This requires a common time reference or base. A global time is an abstract common time reference within a distributed system. It can be derived from synchronizing different clocks in the system to some desired accuracy with respect to a standard of time. The standard of time must be concerned with two aspects: i) it must contain a relative reference in order to measure a duration within a system and ii) it must contain an absolute starting point in order to measure the position of an event on the time line in relation to the environment.

The accuracy of time measurement in a distributed system is limited by the achievable synchronism of the approximate global times in the nodes of the system. Internal and external synchronizations are used to derive a global time base which can properly describe all events and durations in the system.

Clock synchronization algorithms for distributed systems must meet the following requirements:

1. The maximum difference of the time values (correct logical clocks) of any distributed event measured with respect to the approximate global times of any two nodes of the distributed system must be bounded by a known constant δ . This is called *the agreement condition*.

$$|C_i(t) - C_j(t)| \leq \delta \quad \text{for any } i, j, \text{ and } t.$$

2. The deviation of the rate of the approximate global times from the rate of the standard physical time must be limited. This means that correct logical clocks are within a linear envelope of real time. This is called *the accuracy condition*.

$$\frac{t}{(1+\gamma)} + a \leq C_i(t) \leq (1+\gamma)t + b \quad ,$$

for some constants a and b that depend on the initial conditions of the execution of the algorithm, and a constant γ which is a bound on the rate of drift of logical clocks from real time (and hence is a measure of their accuracy with respect to real time). We are interested in

synchronization algorithms that minimize γ . Srikanth and Toueg[87] proved that γ cannot be smaller than ρ , the bound on the accuracy of hardware clocks. That is, $\gamma \geq \rho$. Therefore, for an optimal accuracy algorithm, this condition becomes

$$\frac{t}{(1+\rho)} + a \leq C_i(t) \leq (1+\rho)t + b.$$

3. The clock synchronization algorithm must be fault tolerant, i.e., a fault in any local real time clock and the loss of a message must be tolerated.
4. The execution of the clock synchronization algorithm should not degrade the performance of the system.
5. The synchronization algorithm should be independent of the number of clocks, provided a minimum number of clocks is operational.

The limits of the achievable accuracy are determined by:

1. The variability of the delays of the resynchronization messages.
2. The granularity of the local clock (the period of oscillation or duration between two consecutive ticks)
3. The clock drift during the resynchronization interval which is a measure of the deterioration in the time maintained by a clock between resets.
4. The characteristics of the chosen internal synchronization algorithm which may include reading error, resynchronization intervals, etc.
5. System failures caused by message loss or malicious faults.

We assume that : i) the clocks are initially synchronized to within δ_0 of one another at the starting time $T^{(0)}$. Thus, $|c_p(T^{(0)}) - c_q(T^{(0)})| < \delta_0$ for all processes p and q . ii) if process p is nonfaulty during the real time interval $[t_1, t_2]$, then c_p is a good clock during that interval.

Clock synchrony is maintained by having processes periodically increment their clocks. Incrementing a clock c by K SECONDS means adding K to the value read from the clock; i.e., $c'(T) = c(T+K)$. Clocks are resynchronized every R SECONDS (resynchronization interval). Thus, $T^{(i)} = T^{(0)} + iR$. We define $R^{(i)}$ to be the interval $[T^{(i)}, T^{(i+1)}]$. If c_p represents process p 's physical clock, then resynchronizing the clocks every R SECONDS means having process p use a logical clock $c_p^{(i)}$ on the time interval $R^{(i)}$, with $c_p^{(i)}(T) = c_p(T + K_p^{(i)})$ for some constant $K_p^{(i)}$. It is assumed that $K_p^{(0)} = 0$, so $c_p^{(0)}(T) = c_p(T)$.

A process p is said to be *nonfaulty* up to time $T^{(i+1)}$ if it is nonfaulty during the real-time interval $[c_p^{(0)}(T^{(0)}), c_p^{(i)}(T^{(i+1)})]$. For all processes p and q , if all but at most f processes are nonfaulty up to time $T^{(i+1)}$, then:

$$|c_p^{(i)}(T) - c_q^{(i)}(T)| < \delta \dots\dots(1) \text{ for all } T \text{ in } R^{(i)}; \text{ and } |c_p^{(i+1)} - c_p^{(i)}| < \eta \dots\dots\dots(2)$$

This is called the **clock synchronization condition** which defines the correct synchronization on the interval $R^{(i)}$. Our problem is: find an algorithm for choosing the values $c_p^{(i+1)}$ such that if

the condition above holds for i , then it will hold for $(i+1)$.

So far, we have discussed several issues related to the clock synchronization problem. To complete the subject, there are yet some additional important issues which must be addressed.

1. Synchronization Levels:

One of several advantages to be gained from distributed systems is the high degree of multiprocessing: the system could execute many jobs in parallel. Each of these jobs is usually decomposed into a set of cooperating tasks that communicate with one another during the course of execution. These cooperating tasks are then assigned to *a group* of processors for execution which are often required to be tightly synchronized. A processor group is composed of redundant *clusters*, each of which will execute the same task, i.e. each of the cooperating tasks will be executed on more than one processor in a group, and individual execution results will be voted on. The size of each cluster will depend only on the fault tolerance specification of the system. Obviously, the processors within a cluster need a tighter synchronization among themselves than those in the same group. This leads to three different levels of synchronization:

- A. Tighter intra-cluster (*cluster level*) synchronization: where all processors are operating on the same task.
- B. Tight intra-group (*group level*) synchronization: where all processors are either operating on the same job or on cooperating tasks which are tightly synchronized with respect to each other.
- C. Relatively loose inter-group (*system level*) synchronization: where any two processors belonging to two different groups, and thus operating on unrelated tasks, are loosely synchronized with respect to each other.

Notice that the size of each group would be usually dependent on i) the number of cooperating tasks that are running simultaneously in that group, ii) the degree of redundancy required for reliability.

2. Synchronization Implementation:

We distinguish between instantaneous and continuous synchronization. In instantaneous synchronization, clock's correction is applied to the local time of node j at an instant of time. In continuous synchronization the rate of the local clock is readjusted for the duration of the synchronization interval such that the correction is spread out over the synchronization interval.

Given an algorithm for instantaneous synchronization of the clocks and a bound on how closely synchronized it keeps them, it is easy to devise an algorithm that spreads out the change and to deduce how well it keeps the clocks synchronized.

3. Synchronization Types:

In a distributed system we may distinguish between two kinds of clock synchronization:

- A. **Internal synchronization:** which refers to the construction of an "approximate" global time base among the ensemble of nodes of the distributed system. This is done via a

periodic internal synchronization of the clocks in order to reduce the deviation of good clocks of an ensemble from each other. Internal clock synchronization keeps processor clocks within some maximum relative deviation of each other. It enables one to measure the duration of distributed activities that start on one processor and terminate on another processor, and to totally order distributed events in a manner that closely approximates their real time precedence.

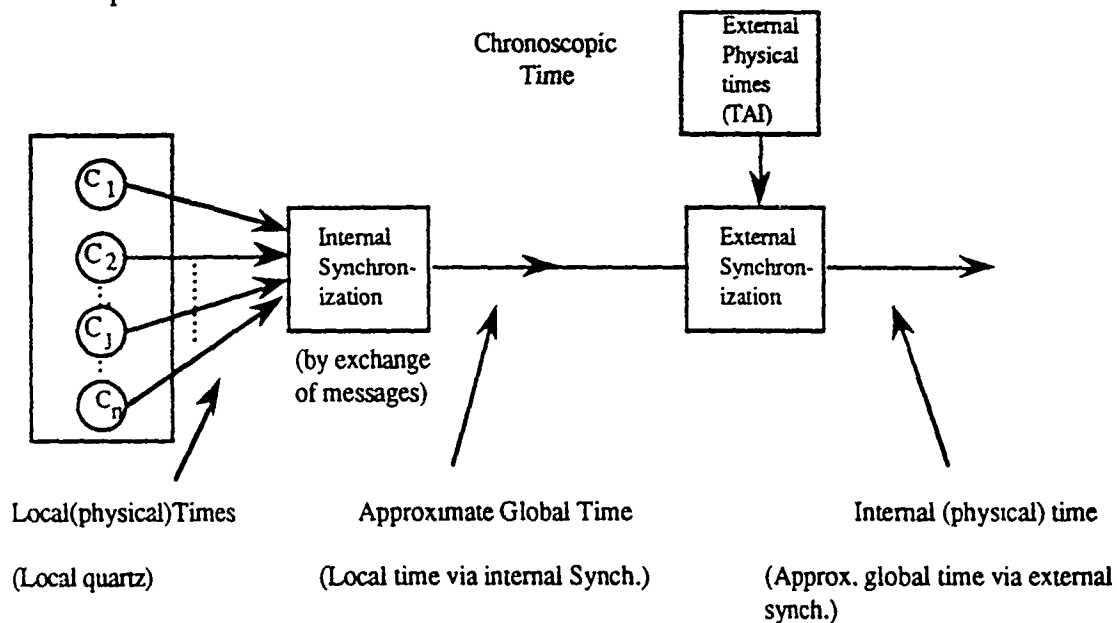


Fig.3. Types of synchronization and the generation of a global time base

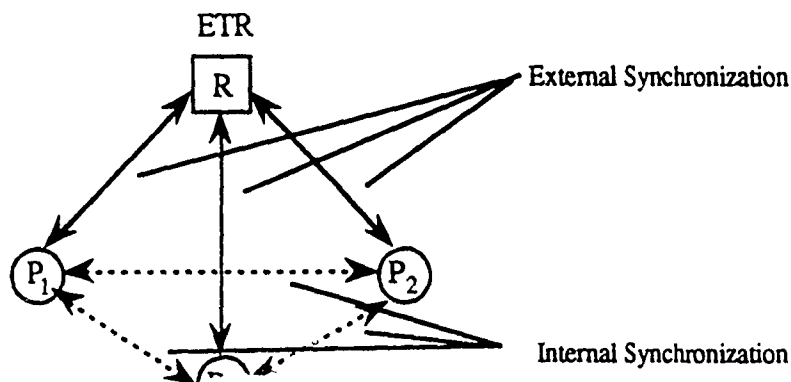


Fig. 4. External and Internal Clock Synchronization

B. External Synchronization: which refers to the synchronization of the approximate global time base mentioned before with the external physical time (real time), in order to generate an internal physical time (see the figure) to act as a common time base. External clock synchronization consists of maintaining processor clocks within some given maximum deviation from a time reference external to the system. Many distributed systems require external clock synchronization to allow exchange of information about the timing of events

with other systems and users, such as ordering the occurrence of related events observed by distinct systems.

Internally synchronized clocks are not necessarily externally synchronized because, as time passes, they can drift arbitrarily far from external time. However, externally synchronized clocks are also internally synchronized.

4. Synchronization Function:

The problem of a set of processes keeping their clocks synchronized can be simply characterized as each process i independently performing a calculation across a distributed set of data:

$$C_i(t) \leftarrow F \{C_{i1}(t), C_{i2}(t), \dots, C_{ik}(t)\}$$

This breaks the problem into two problems: 1) specifying an appropriate way to collect the distributed data; and 2) specifying a synchronization function F .

It is usually assumed that the distributed data is collected by some simple method such as directed broadcasting. Therefore, the works would concentrate on the second problem, namely, the synchronization function. This is a mapping of the data collected about other clocks into a given logical clock.

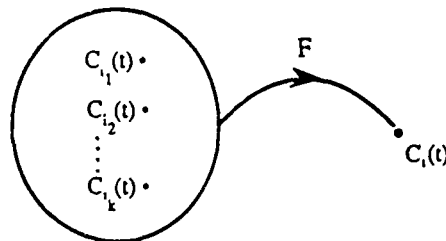


Fig.5. A Synchronization Function

Several synchronization functions have been specified. A simple function that preserves monotonicity is the *maximum value* of the clocks (Lamport [78a]). The *median* clock value and the *mean* value of the clocks have also been used to specify fault-tolerant clock synchronization algorithms (Lamport [82b]). Marzullo and Owicki [83b] proposed two other synchronization functions. The first one depends on the selection of the clock of the time with the *smallest maximum error*. This is possible since each clock module reports its maximum error as required by the algorithm. The second synchronization function is based on the *intersection of the intervals* which is determined by the modules with the latest trailing edge and earliest leading edge. This is a valid function because the correct time must lie in the interval defined to be the intersection of all of the individual intervals.

5. Synchronization Parameters:

The following parameters may be identified as associated with any synchronization procedure.

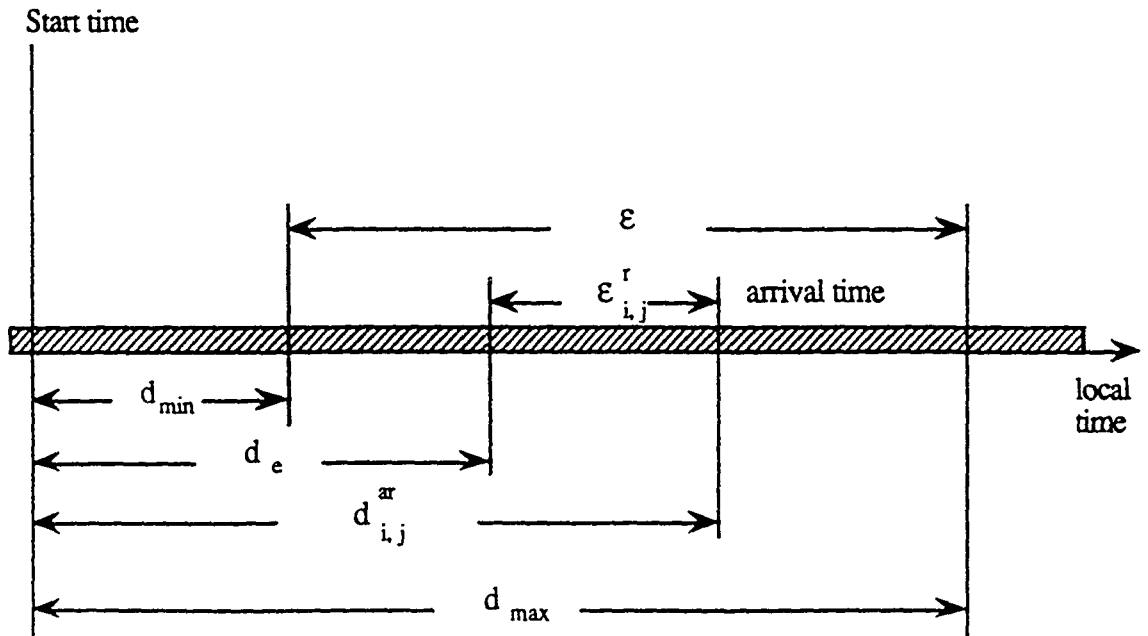


Figure 6.

d_{\min} = minimum delay

d_{\max} = maximum delay

$\epsilon = d_{\max} - d_{\min}$, the reading error.

$\Delta_{i,j}^r$ = the difference between clock i and clock j.

d_e = estimated delay of the message, independent of r.

$d_{i,j}^{ar}$ = actual delay of message from node i to node j in period r.

$\epsilon_{i,j}^r$ = reading delay: a variable delay of message from node i to node j in period r (delay of clock i by node j relative to an estimated message delay).

$x_{i,j}^r$ = the time value of arrival at node j of the synchronization message of clock i, measured in the local time of clock j, with r being a particular synchronization period

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\begin{aligned} x_{i,j}^r &= x_{i,i}^r + (d_e + \epsilon_{i,j}^r + \Delta_{i,j}^r)(1 - \delta_{i,j}), \quad y_{i,j}^r = x_{i,j}^r - (x_{i,i}^r + d_e(1 - \delta_{i,j})) \quad \text{The correction term} \\ &= x_{i,i}^r + (d_{i,j}^{ar} + \Delta_{i,j}^r)(1 - \delta_{i,j}) \quad = (\Delta_{i,j}^r + \epsilon_{i,j}^r) \cdot (1 - \delta_{i,j}) \end{aligned}$$

The accuracy of synchronization is determined by the factors mentioned before (see p.7). The reading error ϵ is determined by the message delays (see that section for details).

6. Synchronization Application:

There are several possibilities for the application of synchronization of clocks. We can distinguish between three alternatives for the analysis of the reading error:

1. We may assume a layered communication system and no control over retransmission or physical access to the communication medium, i.e. clock synchronization is performed in the application software.
2. We may assume that clock synchronization is performed in the kernel of the operating system without special hardware assistance. The send time is written into the message immediately

before transmitting the message. No retransmission of synchronization messages is permitted.

3. We may assume special hardware assistance by a clock synchronization unit. The send time of a message is inserted into the message after access to the communication channel has been granted.

7. Message Delays:

In a distributed real time system, the variable delays of the resynchronization messages are the most significant factors for the accuracy of synchronization. To synchronize processors' clocks (i.e. software synchronization), processes communicate among themselves by sending messages via a communication network. There exists an unpredictable (random) communication delay between the moment a process P_i sends a message to a process P_j and the moment P_j receives the message. This delay corresponds to the time needed to prepare, transmit, and receive a message.

A minimum value of this delay, d_{min} , exists which is the delay measured in the absence of transmission errors and any other system load. On the other hand, an upper bound on message transmission delays, d_{max} , is not known, in general. These delays depend on: 1) the amount of communication and computation going on in parallel in the system, 2) the possibility that transmission errors cause messages to be retransmitted several times, 3) other random events such as page faults, process switches, routes changing, or process freezing by a user.

The general assumption made here is that a message sent from a nonfaulty processor P_i to a nonfaulty processor P_j over a nonfaulty communication link arrives at P_j and is processed at P_j in less than L_d (link delay) real time units. If a message sent from P_i to P_j needs more than L_d to arrive at P_j , or never arrives, then at least one of the processors P_i , P_j or the link is faulty.

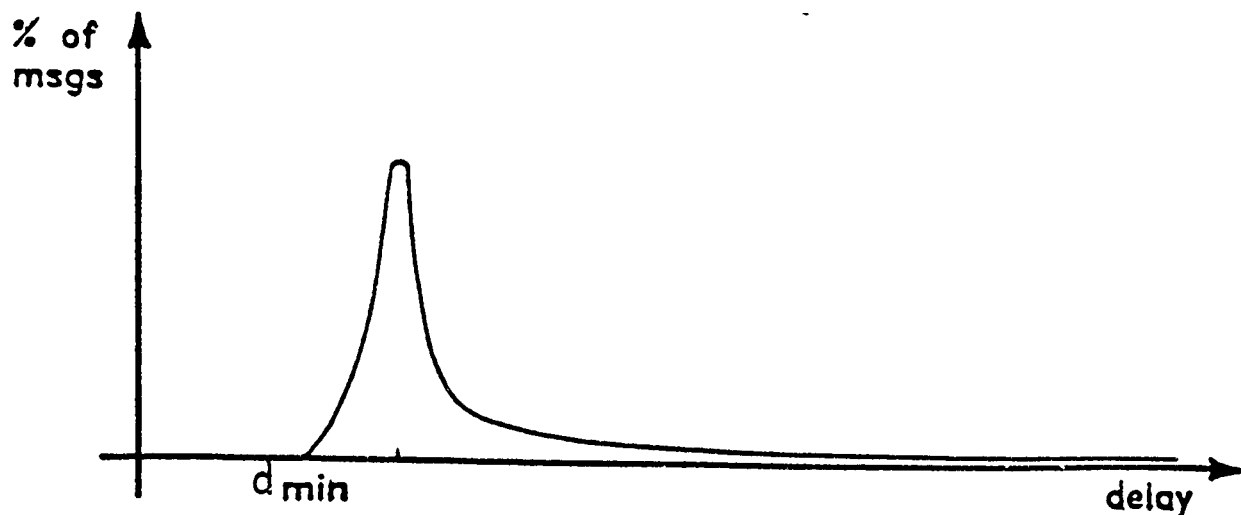


Fig. 7. Distribution of the message delays

Measurement of the process-to-process message delays in existing systems indicate that typically their distribution has a shape resembling that shown in the figure below (Cristian [89a]). This distribution has a maximum density at a mode point between the minimum delay d_{min} and the median delay (usually close to d_{min}), with a long thin tail to the right. See figure 7.

There are two conditions attached to synchronization messages: agreement and validity. Message agreement occurs when each module has an identical copy of that message, that is, all modules agree on the input value. Message validity occurs when all modules have a valid or correct value of that input. Agreement does not imply validity.

In the case of the Byzantine Generals' Problem, the solution proposed identifies the necessary conditions for message agreement in the presence of f arbitrary faults. These conditions are:

- (1) the system must consist of $(3f + 1)$ modules;
- (2) the modules must be interconnected by $(2f + 1)$ disjoint paths;
- (3) the messages must be exchanged $(f + 1)$ times between the participants, and
- (4) the modules must be synchronized to provide a bounded skew.

Messages between clock modules may be of two kinds:

- a) Oral message: one whose contents are completely under the control of the sender. Thus, a faulty sender can transmit any possible message.
- b) Signed message: where anyone can verify the authenticity of a module's signature. Also, it is assumed that a correct module's signature cannot be forged, and any alteration of the contents of its signed messages can be detected. No assumptions about a faulty module's signature are made. It is important to notice here that under signed messages assumption a solution exists that copes with f faults for any number of modules N , with $N \geq f + 2$.

FAULTS AND CLOCK SYNCHRONIZATION:

The issue of faults in a computing system has been studied at length, and many classifications have been introduced. Here we are concerned with faults associated with clocks and clock synchronization. Types of clock failures may be classified as follows:

- a. **Crash failures:** where processes fail by just stopping. Examples of these types of failures: a clock stops, a link breaks down.
- b. **Omission failures:** failures that are due to faults that cause a component never to respond to the occurrence of a trigger event. (i.e. a process occasionally omits to receive, send, or send and receive certain messages).
- c. **Timing failures:** failures that are due to faults that cause a component to respond either too early or too late, that is, outside the interval specified. A change in the oscillator causes the clock value to be incremented too fast or too slowly. Other examples of timing failures may be a too slow process to respond to trigger events (such as message arrivals or timeout occurrences) due to excessive load or slow timers, or a timeout too early because of timers running at speeds greater than $(1+p)$. Timing faults are also referred to as *performance faults*.
- d. **Byzantine failures:** failures that are due to faults that cause a component to respond in a manner different from the one specified. (ex: a non-monotonic time displayed due to some of the bits being stuck at 0 or 1).

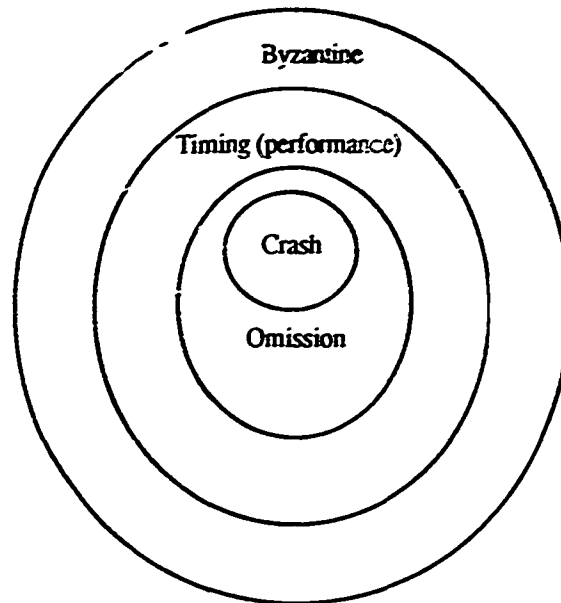


Fig. 8. Types of Clock Faults

Faults are not as arbitrary (Byzantine type) as we may assume. Fault-tolerant clock synchronization algorithms have been studied under simpler, more restrictive models of failures (See for example Hadzilacos [83], Lamport and Fischer [82a], Perry and Toueg [86], Srikanth and Toueg [87] and Thambidurai and Park [88b]). In such cases, developing algorithms for the simpler model of failure could result in easier and less expensive solutions. As an attempt to do just that, the set of all possible faults, F , can be partitioned into two disjoint subsets: Malicious faults M , and non-malicious faults B .

A *malicious fault* refers to the case when not every non-faulty receiver can detect that a fault has occurred. It is these faults that require the use of Byzantine Agreement algorithms with several rounds of message exchange. Maliciously faulty processors send different time information about themselves to different nonfaulty processors during the course of synchronization. A *non-malicious fault* (also called intercepted fault) is defined to be a fault which is detected by every non-faulty receiver of the message. The detection must occur prior to the use of the message in the algorithm. Examples of non-malicious faults include: timing faults (Cristian et al. [85], Meyer et al. [87]), omission faults (Cristian et al. [85], Srikanth et al. [87]), crash faults (Babaoglu et al. [85]).

Each of the two subsets above can be partitioned further into two disjoint subsets: symmetric faults S , and asymmetric faults A (Thambidurai et al. [88b]). A *symmetric fault* refers to the case when all receivers obtain exactly the same message. Symmetric faults may occur as a result of faulty processors or as a result of faulty communication channels (links). *Asymmetric faults* refer to the case when a message is not received identically by all the non-faulty receivers of that message, i.e. at least one non-faulty recipient receives a different message.

According to Thambidurai et al. [88b], the probability of an asymmetric fault could be much

less than the probability of a symmetric fault. Likewise, non-malicious faults are far more likely to occur than malicious faults. That is, in practice:

$$Pr \{ \text{Non-malicious fault} \} > Pr \{ \text{malicious symmetric fault} \} > Pr \{ \text{malicious Asymmetric fault} \}$$

If only non-malicious faults are considered, the algorithms for agreement are simple and efficient. For example, it was shown in the above work and in Vasanthavada et al. [89] that it is possible to guarantee Byzantine Agreement with much less processor and message overhead when faults are partitioned into non-malicious faults, symmetric faults, and asymmetric faults.

FAULT-TOLERANT CLOCKS:

Clock synchronization would not have been much of a problem had all the clocks, including the faulty ones, behaved consistently with one another. However, when some of the faulty clocks behave in any arbitrary manner, clock synchronization can become a serious problem. A faulty clock can make it difficult for the other clocks to synchronize themselves by either omitting or sending conflicting clock information during the course of synchronization. This kind of arbitrary behavior is referred to as a Byzantine fault [Lamport et al. 85]. We would like an algorithm in which a fault in one of the processes or in its clock does not prevent the other processes from synchronizing their clocks. If at most f clocking modules fail (f being the degree of fault tolerance), the remaining modules will still produce synchronous clock signals.

A fault-tolerant clock is not a single clocking source but is independently derived in each channel by a majority vote of a redundant set of clocks. In the presence of f malicious faults, $(3f + 1)$ clocks rather than the simple majority voting scheme that uses $(2f + 1)$ clocks is required to design a fault-tolerant clock.

Loss of synchronization among individual clock modules can be caused by failure modes affecting the clock frequency, phase, and amplitude. Here are some important types of clock failures that may be encountered:

1. "Stuck-at" Type Faults:

Stuck-at 0/1, and stuck-open faults lead to time invariant and hence fixed phase behavior of the faulty modules as perceived by the clock receivers of the other modules. If clock receivers are transition sensitive rather than level sensitive, the effect of the faulty clock will be minimized. As a result of the transition sensitive design, such faults may have no impact on the synchrony of the non-faulty modules in steady-state.

2. The Multiple Cliques Problem:

Multiple cliques are self-sustaining groups of clocks in which the non-faulty clocks maintain synchrony within their own groups, but are out of synchrony with the other non-faulty clocks in the system. This problem can occur in the presence of malicious failures of the faulty clocks.

3. Byzantine Faults:

A malfunctioning component may exhibit the following type of behavior: it sends conflicting

information to different parts of the system. The problem of coping with this type of failure is expressed abstractly as the Byzantine Generals Problem.

Several approaches have appeared in the literature as solutions to this problem. The following results have been proved, which are just as good for fault-tolerant system of clocks with f arbitrary (or Byzantine) faults:

- 1) The system must consist of at least $(3f + 1)$ modules;
- 2) The modules must be interconnected through $(2f + 1)$ disjoint paths;
- 3) The information must be exchanged $(f + 1)$ times between the participants.

Reaching an approximate agreement is just as hard and costly as reaching exact agreement (see Lamport et al. [82c]).

CLOCK SYNCHRONIZATION SCHEMES

We have presented so far, very briefly, some basic theoretical background and issues defining the problem of clock synchronization in a distributed computing system. This problem has been studied extensively in recent years, and a number of different synchronization schemes have been proposed. See figures 9 -13.

Hardware schemes provide tight synchronization through the use of special hardware. Separate network of clocks may be required. However, minimal time overheads on systems performance can be achieved.

Software schemes use special synchronization messages to synchronize clocks, thus requiring no special hardware. However, these schemes do not provide tight synchronization, yet they are flexible and can be changed as desired. Probabilistic (software) schemes provide tight synchronization and use no special hardware, but require greater number of synchronization messages which may make it not suitable for large distributed systems.

Hybrid models, using a combination of hardware and software components, provide a balance between the hardware requirements and the clock skew attainable for a given number of synchronization messages. Thus, lesser hardware is required (no separate network of clocks), and tighter bounds than software solutions can be achieved.

There have been several approaches proposed in the literature to solve the problem of clock synchronization using one of the techniques described. These attempts can be classified into three main categories as mentioned above and shown in figure 9.

We have reviewed extensively most of the literature written on this problem using hardware solutions, probabilistic software solutions, and hybrid solutions. We did review some of the work done using deterministic software algorithms. However, there is much more to be reviewed in this area. The figures below summarize as well as categorize what we reviewed and covered so far. For the details of the review, please refer to the technical report.

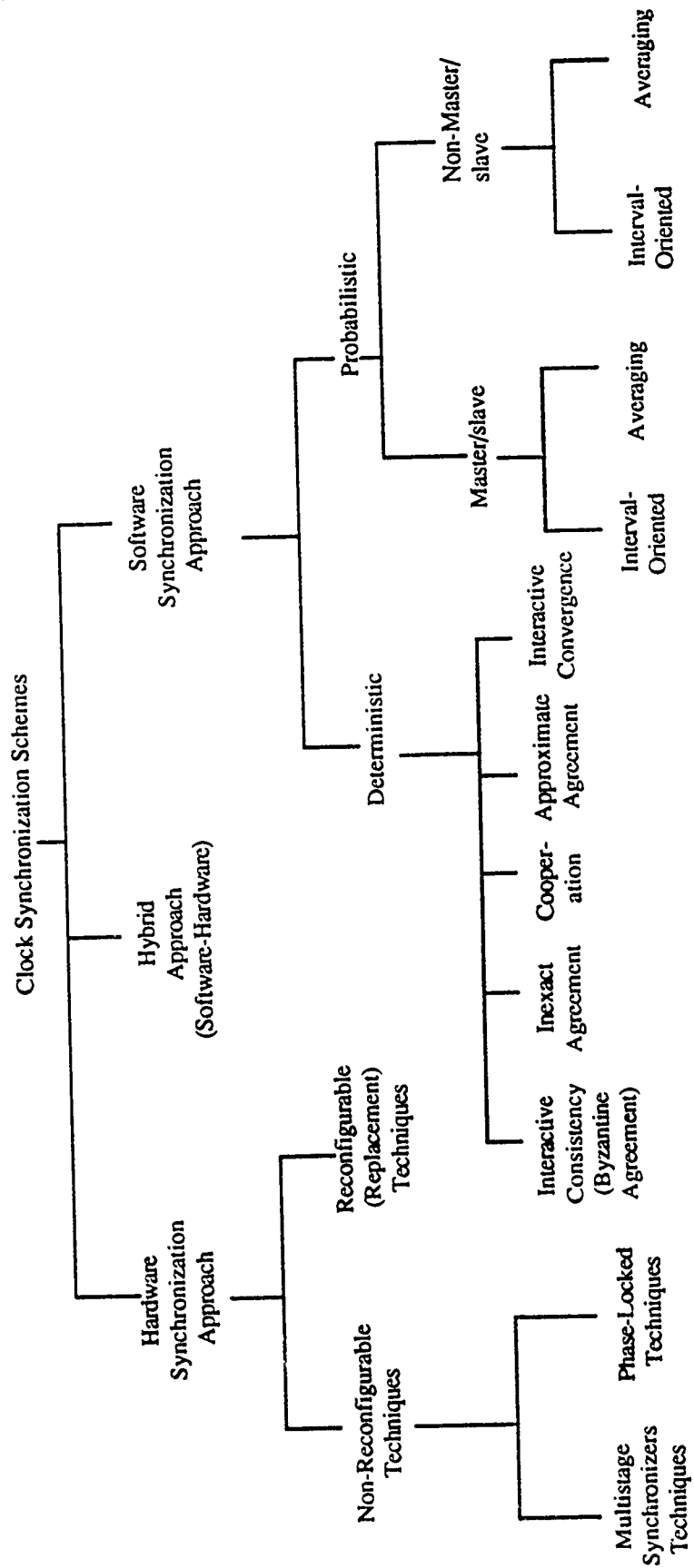


Fig. 9 Clock Synchronization Schemes

1. Hardware Synchronization Schemes Proposed

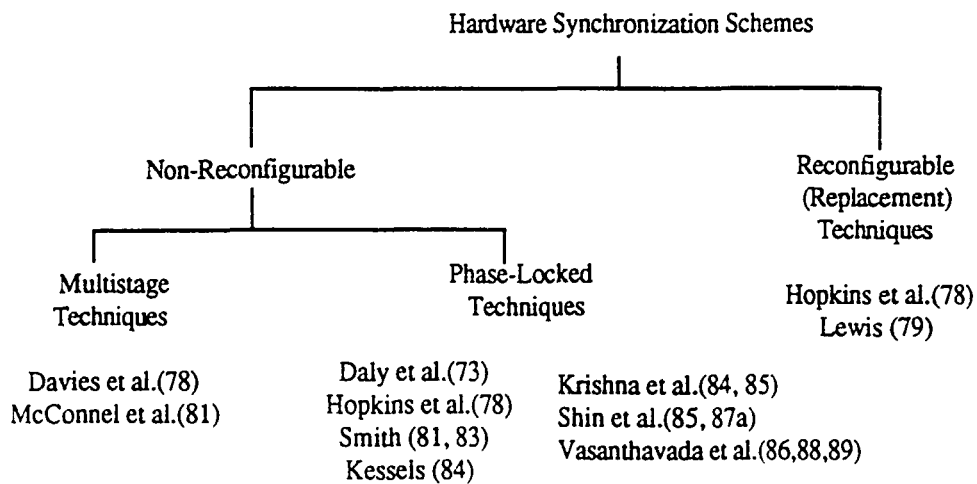


Fig. 10 Hardware techniques proposed

2. Software clock synchronization algorithms proposed:

A. Probabilistic Clock Synchronization Schemes Proposed:

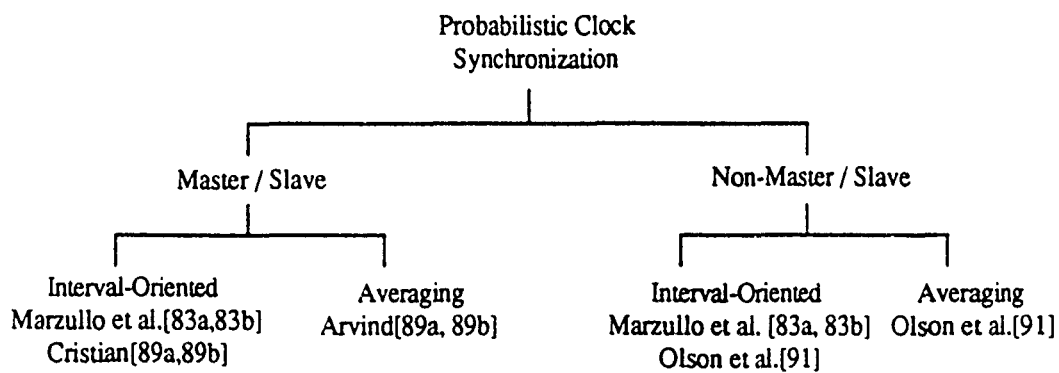


Fig.11 Probabilistic Clock Synchronization Schemes Proposed

B. Deterministic Software Clock Synchronization Schemes Proposed:

Deterministic Software Synchronization Algorithms

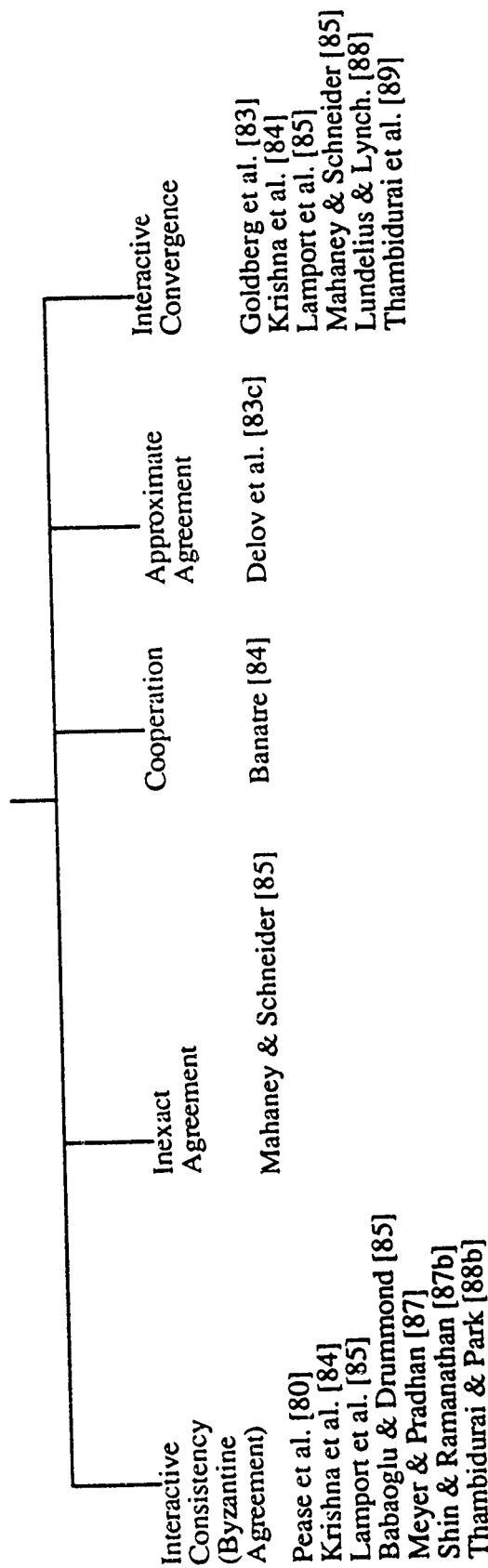


Fig. 12 Deterministic Software Synchronization
Algorithms Proposed

3. Hybrid Clock Synchronization Schemes Proposed:

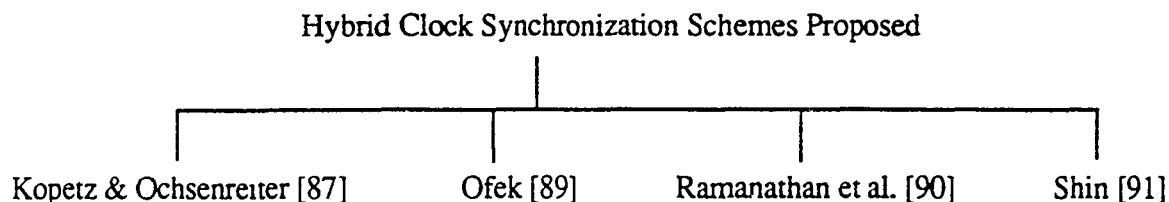


Fig.13 Hybrid schemes proposed

CONCLUSION:

Given a network of N processors, each of which has a clock of its own, the problem is to synchronize all the nonfaulty clocks in the network to a specified fault tolerance using as few interconnections as possible in such a way that symmetry of the network is retained.

Synchronized clocks are useful for a number of reasons: 1) They can be used to totally order the events of a distributed system in a manner which is fairly consistent with real-time event precedence, e.g. merging of the data logs generated on distinct computers into a common unique log. 2) They can be used to measure, with an a priori known accuracy, the time that elapsed between events that occur on different processors, e.g., do performance evaluation of distributed systems. 3) They can be used to insure consistency among the knowledge states of the computers of a distributed system.

In a fault-tolerant multiprocess system, it is often necessary for the individual processes to maintain clocks that are synchronized with one another. Physical clocks do not keep perfect time, rather they drift with respect to one another. Therefore, the clocks must periodically be resynchronized. Such a fault-tolerant system needs a clock synchronization algorithm that works despite faulty behavior by some processes and clocks. Implementing reliable clock synchronization presents many problems, some of which are: achieving initial synchrony, building accurate hardware clocks, designing programming language primitives for reading the clocks, and restarting a failed process and bringing it into synchrony with the other processes.

In this report we have attempted to identify and summarize most of the issues associated with the clock synchronization problem in a distributed computing system. This was done in order to define this problem's domain and address the issues involved in the solution. As this discussion indicated, the problem is a complicated one which explains why there has been so much written on it. Nevertheless, there is still a need for further studies of this problem, particularly on the subjects of evaluation and comparison of the existing schemes, as well as on developing new, reliable and more efficient ones. We hope that it will be possible to complete this review by covering what is left on deterministic clock synchronization schemes in the near future.

ACKNOWLEDGEMENTS

The author wishes to thank Richard Metzger and Raymond Liuzzi for their support, Jon Valente and Robert Vaeth for the helpful discussions and ideas, and Rex Gantenbein for the productive time we had together. This work was supported by the AFOSR Summer Faculty Research Program. The views and opinions expressed in this report are those of the author.

REFERENCES

- Arvind, K., (89a); "A New Probabilistic Algorithm for Clock Synchronization," Proc. of Real-Time Systems Symp., pp.330-339, Santa Monica, CA.
- Arvind, K., (89b); "A New Probabilistic Algorithm for Clock Synchronization," COINS Technical Report 89-86, Department of Computer and Information Science, University of Massachusetts at Amherst, Amherst, Massachusetts, August 1989. *
- Babaoglu, O., and R. Drummond, (85); "Streets of Byzantium: Network Architectures for Fast Reliable Broadcasts," IEEE Trans. on Software Eng., Vol. SE-11, No. 6, June 1985, pp.546-554.
- Banatre, J.P., and M. Banatre, (84); "A Cooperation Scheme and its applications to Clock Resynchronization in Distributed Systems," Parallel Computing 83, M. Feilmeier, J. Joubert, and U. Schendel (eds.), Elsevier Science Publishers, 1984, pp. 381-389.
- Cristian, F., (89a); "A Probabilistic Approach to Distributed Clock Synchronization," Proc. of the Ninth International Conference on Distributed Computing Systems, May 1989, pp.288-296.
- Cristian, F., (89b); "Probabilistic Clock Synchronization," Distributed Computing, Vol. 3, pp. 146-158, 1989. *
- Cristian, F., H. Aghili, R. Strong, and D. Dolev, (85); "Atomic Broadcast: from Simple Message Diffusion to Byzantine Agreement," Proc. 15th Intern. Symp. on Fault-Tolerant Computing, IEEE Press, pp.200-206.
- Daly, W.M., A.L. Hopkins, and J.F. McKenna, (73); "A Fault-Tolerant Digital Clocking System," Dig. 3rd Int. Symp. of Fault-Tolerant Comput., June 1973, pp.17-22.
- Davies, D., and J.F. Wakerly, (78); "Synchronization and Matching in Redundant Systems," IEEE Trans. on Computers, Vol.27, No.6, June 1978, pp.531-539.
- Dolev, D., N.A. Lynch, S.S. Pinter, E.W. Stark, and W.E. Weihl, (83c/86), "Reaching Approximate Agreement in the Presence of Faults," Proc. Third Symp. on Reliability Distributed Software Database Syst., Oct. 1983, pp. 145-154. Also in JACM, Vol. 33, No. 3, July 1986, pp. 499-516.
- Goldberg, J., et al., (83); "Development and Analysis of the Software Implemented Fault-Tolerance (SIFT) Computer," NASA Contract Report CR-172146, June 1983. *
- Hadzilacos, V., (83); "Byzantine Agreement Under Restricted Types of Failures (not telling the truth is different from telling lies)," Tech. Rep. 19-83, Aiken Computation Laboratory, Harvard Univ., Cambridge, Mass., June 1983. *
- Halpern, J. Y., B. Simons, R. Strong, D. and D. Dolev, (84); "Fault-Tolerant Clock Synchronization," Proc. of the Third Annual ACM Symposium on Principles of Distributed Computing, pp.89-102, August 1984, Vancouver, Canada. *
- Hopkins, A.L., T.B. Smith, and J.H. Lala, (78); "FTMP-A Highly Reliable Fault-Tolerant Multiprocessor for Aircraft," Proc. of the IEEE, Vol.66, No.10, Oct.1978, pp.1231-1239.

- Kessels, J.L.W., (84); "*Two Designs of a Fault-Tolerant Clocking System*," IEEE Trans. Computers, Vol. C-33, No.10, pp.912-919, Oct. 1984.
- Kopetz, H., (86); "*Accuracy of Time Measurement in Distributed Real Time Systems*," Proc. of Distributed Software DB Systems, IEEE Press, Los Angeles, CA, 1986, pp.35-41.
- Kopetz, H., and W. Ochsenreiter, (87); "*Clock Synchronization in Distributed Real-Time Systems*," IEEE Transactions on Computers, Vol. C-36, No.8, pp.933-940, August 1987.
- Krishna, C.M., K.G. Shin, and R.W. Butler, (85); "*Ensuring Fault-Tolerance of Phase-Locked Clocks*," IEEE Trans. Computers, Vol. C-34, No.8, pp.752-756, Aug. 1985.
- Krishna, C.M., K.G. Shin, and R.W. Butler, (84); "*Synchronization and Fault-Masking in Redundant Real-Time Systems*," Proc. 14th Intern. Symp. on Fault-Tolerant Computing, June 1984, pp.152-157.
- Lamport, L., (78a); "*Time, Clocks, and the Ordering of Events in a Distributed System*," Communication of the ACM, Vol.21, No.7, July 1978, pp.558-565.
- Lamport, L., and M. Fischer, (82a); "*Byzantine Generals and Transaction Commit Protocols*," Opus 62, SRI International, Menlo Park, Calif., Apr. 1982. *
- Lamport, L., P.M. Melliar-Smith, (82b); "*Synchronizing Clocks in the Presence of Faults*," CSL (Lamport Opus 60) Tech. Rep. 141, SRI International, 1982. *
- Lamport, L., R. Shostak, and M. Pease, (82c); "*The Byzantine Generals Problem*," ACM Trans. Programming Languages & Syst., Vol.4, No.3, pp.382-401, July 1982.
- Lamport, L., P.M. Melliar-Smith, (85); "*Synchronizing Clocks in the Presence of Faults*," Journal of the Association for Computing Machinery, Vol.32, No.1, January 1985, pp.52-78.
- Lewis, D.W., (79); "*A Fault-Tolerant Clock Using Standby Sparing*," Proc. of the 9th Intern. Symp. on Fault-Tolerant Computing, IEEE Press, pp.33-39.
- Lundelius-Welch, J., and N. Lynch, (88); "*A New Fault-Tolerant Algorithm for Clock Synchronization*," Information and Computation, Vol.77, No.1, pp.1-36, 1988.
- Mahaney, S.R., and F.B. Schneider, (85); "*Inexact Agreement: Accuracy, Precision and Graceful Degradation*," Proceedings of the 4th Symposium on the Principles of Distributed Computing (Minaki, Canada), ACM, New York, Aug. 1985, pp.237-249.
- Marzullo, K., (83a); "*Maintaining the Time in a Distributed System. An Example of a Loosely-Coupled Distributed Service*," Ph.D. Dissertation, Dept. of Electrical Engineering, Stanford Univ., Stanford, Calif., 1983. *
- Marzullo, K. and S.S. Owicki, (83b); "*Maintaining the Time in a Distributed System*," Proc. of the 2nd ACM SIGACT-SIGOPS Symp. on Principles of Distributed Computing, August 1983, Montreal, Canada, ACM, pp.295-305.
- McConnel, S.R. and D.P. Siewiorek, (81); "*Synchronization and Voting*," IEEE Trans. on Computers, Vol.30, No.2, Feb.1981, pp.161-164.
- Meyer, F.J., and D.K. Pradhan, (87), "*Consensus with Dual Failure Modes*," Proc. of the 17th Intern. Symp. on Fault-Tolerant Computing, July 1987, pp.48-54.
- Ofek, Y., (89); "*Generating a Fault-Tolerant Global Clock in a High Speed Distributed System*," Proc. of the 9th Intern. Conf. on Distributed Computing Systems, IEEE Press, 1989, pp.218-226.
- Olson, A. and K.G. Shin, (91); "*Probabilistic Clock Synchronization in Large Distributed Systems*," Proc. of the

11th Intern. Conf. on Distributed Computing Systems, May 1991, IEEE Press.

Pease, M., R. Shostak, and L. Lamport, (80); "Reaching Agreement in the Presence of Faults," J. of ACM, Vol.27, No.2, April 1980, pp.228-234.

Perry, K.J., and S. Toueg, (86); "Distributed Agreement in the Presence of Processor and Communication Faults," IEEE Trans. Software Eng., SE-12, No. 3, Mar.1986, pp. 477-482.

Ramanathan, P., D.D. Kandlur, and K.G. Shin, (90); "Hardware-Assisted Software Clock Synchronization for Homogeneous Distributed Systems," IEEE Trans. on Computers, Vol.39, No.4, pp.514-524, Apr. 1990.

Shin, K.G., (91); "HARTS: A Distributed Real-Time Architecture," Computer Magazine, Vol. 24, No.5, May 1991, pp.25-35.

Shin, K.G. and P. Ramanathan, (87a); "Clock Synchronization of a Large Multiprocessor System in the Presence of Malicious Faults," IEEE Trans. Computers, Vol. C-36, No.1, pp.2-12, Jan. 1987.

Shin, K.G. and P. Ramanathan, (87b); "Diagnosis of Processors with Byzantine Faults in a Distributed Computing System," Proc. of the 17th Intern. Symp. on Fault-Tolerant Computing, IEEE Press, 1987, pp.55-60.

Shin, K.G. and P. Ramanathan, (85); "Synchronization of a Large Clock Network in the Presence of Malicious Faults," Proc. of the Real-Time Systems Symp., IEEE Press, pp.13-24, Dec.1985.

Smith, T.B., (81); "Fault-Tolerant Clocking System," Dig. 11th Intern. Symp. Fault-Tolerant Comput., June 1981, pp.262-264.

Smith, T.B., and J.H. Lala, (83); "Development and Evaluation of a Fault-Tolerant Multiprocessor (FTMP) Computer. Volume I: FTMP Principles of Operations," NASA Contractor Report CR-166071, May 1983. *

Srikanth, T.K. and S. Toueg, (85/87); "Optimal Clock Synchronization," J. ACM, Vol.34, No.3, pp.626-645, July 1987. Also in Proc. of the 4th Annual ACM Symp. on the Principles of Distributed Computing, ACM, 1985, pp.71-86.

Srikanth, T.K. and S. Toueg, (84/87); "Simulating Authenticated Broadcasts to Derive Simple Fault-Tolerant Algorithms," Tech. Rep. 84-623, Dept. of Computer Science, Cornell Univ., Ithaca, N.Y., July 1984. Also in Distributed Computing, Vol.2, pp.80-94, 1987, Springer-Verlag, New York. *

Thambidurai, P., and Y.K. Park, (88b); "Interactive Consistency with Multiple Failure Modes," Proc. of the 7th Symp. on Reliable Distributed Systems, IEEE Press, 1988, pp.93-100.

Thambidurai, P., Y.K. Park, and K.S. Trivedi, (89b); "On Reliability Modeling of Fault-Tolerant Distributed Systems," 9th Intern. Conf. on Distributed Computing Systems, pp. 136-142, 1989.

Thambidurai, P., A.M. Finn, R.M. Kieckhafer, and C.J. Walter, (89a); "Clock Synchronization in MAFT," Proc. of the 19th Intern. Symp. on Fault-Tolerant Computing, IEEE Press, 1989, pp.142-149.

Vasanthavada, N., P.N. Marinos, and G.S. Mersten, (86b); "Design and Performance Evaluation of Mutually Synchronized Fault-Tolerant Clock Systems," Dig. 16th Intern. Symp. Fault-Tolerant Comput., July 1986, pp.206-211.

Vasanthavada, N., P.N. Marinos, (88); "Synchronization of Fault-Tolerant Clocks in the Presence of Malicious Failures," IEEE Trans. Computers, Vol.37, No.4, pp.440-448, Apr. 1988.

Vasanthavada, N., P. Thambidurai, and P.N. Marinos, (89); "Design of Fault-Tolerant Clocks with Realistic Failure Assumptions," Proc. 19th Intern. Symp. on Fault-Tolerant Computing, IEEE Press, 1989, pp.128-133.

(References marked with * were inaccessible but their major results have been reported in other references).

CHARACTERIZATION OF RADAR CLUTTER AS AN SIRP

CHARLES T. WIDENER and JAY K. LEE, Ph.D.

ABSTRACT

It has been proposed that radar clutter can be modeled as a spherically invariant random process or SIRP. SIRPs seem well suited to this role since by variation of certain parameters the Weibull, K- or Rayleigh distributed clutter envelopes are obtained. These distributions are significant since they fit well with experimental radar clutter data under different circumstances. In this report, a radar clutter model based on rough surface scattering is developed to show that SIRP characterization can be based on electromagnetic principles. Small perturbation analysis of a two-scale randomly rough surface is chosen since the form for the backscattered field has the proper form for an SIRP, under certain conditions. This is an important step in being able to predict the proper statistical distribution of radar clutter based on surface geometry and electromagnetic properties.

INTRODUCTION

The characterization of radar clutter has been an important study for radar engineers, from the inception of radar, to the present. Due to the random nature of background terrain, statistical characterization seems appropriate from both theoretical and experimental considerations. Gaussian models have been commonly used, but they become inadequate in many cases of interest. Other characterizations have been proposed on the basis of empirical studies, examples being Rician [1], Weibull [2], log-normal [3], and K-type [4, 5]. More recently it has been proposed that radar clutter be

modeled as a complex random process [6], more specifically as a spherically invariant random process (SIRP) [7-9]. Briefly, an SIRP is such that given a real or complex random process

$$\bar{X}(t) = \{X_1(t), X_2(t), \dots, X_n(t)\}^T$$

every sample taken from it, that is

$$\bar{X} = \{X_1(t_1), X_2(t_2), \dots, X_n(t_n)\}^T$$

is spherically invariant random vector (SIRV) with one and the same characteristic cumulative distribution function (CCDF). An SIRV can be thought of as being obtained by a product of a Gaussian random vector times an independent non-negative random scale factor s with some CDF, which is the CCDF of SIRV under consideration. The usefulness of the SIRP characterization comes by varying parameters of s to obtain different clutter statistics. It can be shown [6] that for certain parameters the Rayleigh, K-, or Weibull clutter envelopes result. This unifying concept is an important step in further understanding clutter.

The purpose of this report is to develop a mathematical model for clutter from electromagnetic theory, suitable as a basis for an SIRP characterization. Specifically, the form of received signal (from clutter) should be of the form $V(t) = G(t) \cdot s$ where $G(t)$ is a Gaussian random process, s is a random variable, and $V(t)$ is the received signal. Without regard to type of receiver, the signal $V(t)$ sought is the backscattered field $\vec{E}^s(t)$. As a further specialization, concentration will be on a single sample (or instance of time) corresponding to the returns from a single radar illumination area or "footprint".

DISCUSSION OF APPROPRIATE MODELS

In the attempt to derive useful, accurate mathematical models for radar clutter,

two approaches are commonly used. The first views the illuminated area or target background as a collection of point scatterers [10]. These scatterers are assumed randomly distributed over a surface, each having an associated amplitude and phase. In the most general case, the amplitude and phase distributions are random variables. Classically, by making various assumptions about these distributions, different clutter characteristics are derived, e.g. a Rayleigh distributed envelope results when the phase is assumed equal and constant. This phenomenological approach is useful, but not really derived from first principles.

A second approach, and the one used in this report, is treating radar clutter as scattering from a rough, randomly distributed surface. A rough surface can be described in a system of Cartesian coordinates as

$$Z = \zeta(x, y)$$

where ζ is a random variable depending on an x and y . The mean level of the surface is the plane

$$Z = 0$$

Any point on the surface can be described by a position vector from the origin

$$\vec{r} = \hat{x}x + \hat{y}y + \hat{z}\zeta(x, y)$$

A random surface can be described by the statistical distribution of its deviation from a mean level. This, however, does not completely describe the surface; it says nothing about how close or far apart the hills and depressions are. A second function, the correlation function, or its normalized version, the correlation coefficient describes this aspect of the surface. The correlation coefficient $C(\tau)$ gives a measure of dependence or

correlation of two points $\zeta(x_1, y_1)$ and $\zeta(x_2, y_2)$ on the surface. If the separation of the points

$$\tau = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

is very large, (not on the same hill or valley) then the points are independent. Conversely, points near one another will be correlated; when $\tau=0$, they are the same point, or

$$\lim_{\tau \rightarrow 0} C(\tau) = 1$$

The distance T in which $C(\tau)$ drops to e^{-1} is called the correlation length.

A commonly used distribution is the normal distribution with zero mean

$$w(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$$

where σ is the rms value of the deviation of ζ from $z=0$. A commonly used correlation coefficient is

$$C(\tau) = \exp\left(-\tau^2/T^2\right)$$

Specification of σ and T makes the rough surface model approximate a wide variety of rough surfaces met in practice.

Scattering of waves (acoustic or electromagnetic) from a rough surface is a complex problem studied by many people in this century and yet it remains a popular topic for research papers. Many methods of analysis have been used to study scattering from a rough surface, but usually each method assumes a priori some condition or scale limitation to make the problem tractable.

Two important examples of the methods are the physical optics methods (or tangent plane approximation) and the small perturbation method (SPM). In the former an assumption is made that the incident wavelength is much smaller than the variations of the surface. Under this assumption, an

incident wave locally "sees" a planar surface. Reflection (and transmission) at a planar interface is a classic problem in electromagnetics, so by integrating over the surface, a solution can be found. Mathematically, the validity of the solution is restricted by the condition, derived by Brekhovskikh [11], that the radius of curvature of the surface is much greater than the incident wavelength or

$$4\pi r_c \cos \vartheta \gg \lambda \quad (\text{as reported by Beckmann})$$

where r_c is the minimum radius of curvature in any direction and ϑ is the angle between the incident propagation vector \vec{k} and the local normal \vec{n} . A second condition [10], which intuitively needs to be satisfied, given the condition on r_c is; the correlation length must exceed the wavelength

$$T > \lambda$$

(Beckmann's formulation [see 10] is classic and widely used.)

Sometimes a condition is used for stationary-phase or geometric optics, which includes that of physical optics, [12, 13] namely

$$4k^2 \cos^2 \theta \sigma^2 \gg 1$$

This has the advantage that the rms height deviation σ is related directly to $\lambda_1 (k = \frac{2\pi}{\lambda})$. Because the rms height σ and correlation length T are large compared to λ , this type of surface is called large scale rough surface.

At the other extreme is the SPM. In this method it is assumed that the rms height is small compared to the wavelength

$$4k^2 \cos^2 \theta \sigma^2 \ll 1$$

where θ is the angle of incidence.

Rice solved this problem using a two-dimensional Fourier series expansion of the surface [14]. The coefficients of the Fourier components are deter-

mined from boundary conditions on the surface. A variation on this method was developed by Burrows [15], where by also assuming the rms slope is small,

$$(\nabla_s \zeta)^2 \ll 1$$

the surface equation can be expanded in a Taylor series about its mean. Using the small slope assumption, higher order terms can be neglected. Then, in a manner similar to Rice, the exact boundary conditions on the perturbed surface rewritten as conditions which apply to the unperturbed planar surface and a solution is found. As may be guessed such surfaces are called small scale.

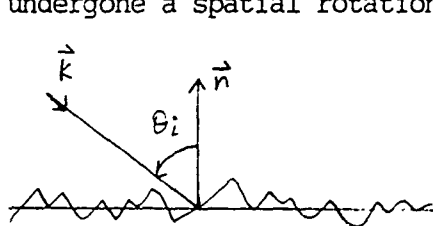
When modeling surfaces occurring in nature, the assumption of exclusively large or small scale is restrictive and unrealistic when compared to modern radar wavelength. If, for example, an ocean were to be modeled as a large scale surface, the radar wavelength would need to be submillimeter due to the presence of fine, wind produced ripples. Conversely, if it were modeled as a small scale surface, the corresponding radar wavelength would be on the order of tens of meters, because of the large sea swells or gravity waves. Such is the case with most surfaces of interest with regard to radar clutter.

To accommodate analysis of more realistic surfaces, a two-scale or composite model has been proposed and studied by several authors [12, 16-21]. A composite surface is simply viewed as a super position of a small scale variation on a large scale variation.

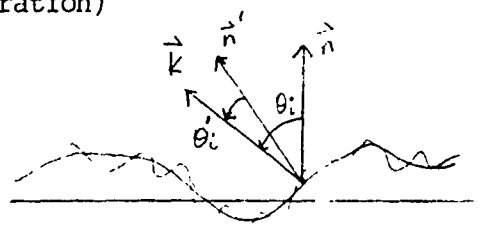
$$\zeta = \zeta_L + \zeta_S$$

The division between large scale and small scale occurs naturally once the radar wavelength is specified. While this model has some limitations of its own, it is far more versatile than either the large or small scale model alone.

Two of the more interesting and simple approaches to dealing with composite surfaces are the tilted plane method and a more generalized SPM. In the tilted plane method, the large scale variation is assumed to present locally flat titled plane areas with superimposed small scale variation. The only difference in the analysis between this and any other small scale model is that the angle of incidence with regard to the normal has effectively undergone a spatial rotation. (see illustration)



a) wave incident upon small scale surface



b) wave incident upon composite surface

The SPM applied to a composite surface is also similar to that applied to an exclusively small scale surface. The exact boundary conditions on the surface are transferred to a system of boundary conditions on the large scale surface instead of the surface $z=0$. Effectively, the variations of scattered field due to the small scale roughness are viewed as being caused by equivalent surface current densities which create the same resultant field as the small scale roughness. The total scattered field is the sum of zeroth order component, due only to the small scale variation. Brown [12] and Bass et al. [19, 20] have both applied this method to a composite scale surface.

Because of the relatively simple form of the scattered field coupled with the fact that cross polarization terms are built in the final form, the SPM for composite surface is useful as a basis to characterize radar clutter as an SIRP. A short summary of this method appears in the following section.

SUMMARY OF SMALL PERTURBATION ANALYSIS

The first case treated is that of a small scale surface only. This is done to illustrate the main points of the analysis without becoming unduly bogged down in cumbersome details. Following this summary, it is shown how small perturbation analysis is applied to a composite surface as a logical next step.

In the scattering analysis of a small scale surface, the surface (denoted Σ) is represented by normal deviation $\zeta(\vec{r})$ from a smooth surface $z=0$ denoted S_0 .



The normal to the surface Σ is denoted \vec{N} , while the normal to the surface $z=0$ is denoted \vec{n} . It is assumed that the rough surface divides a vacuum and medium with arbitrary permittivity ϵ . Also it is assumed that the permeability μ is constant across the interface. Boundary conditions on the surface Σ have the form

$$\begin{aligned} [\vec{N} \times (\vec{E}_1 - \vec{E}_2)]_{\Sigma} &= 0, \quad [\vec{N} \cdot (\vec{E}_1 - \epsilon \vec{E}_2)]_{\Sigma} = 0 \\ (\vec{H}_1 - \vec{H}_2)_{\Sigma} &= 0 \end{aligned} \quad (1)$$

The subscript 1 refers to the fields in the first medium, while the subscript 2 refers to the fields in the second medium. Assuming the normal

deviations ζ and the slopes $\vec{\gamma} = \nabla_S \zeta$ of the surface Σ relative to the surface S_0 are small

$$(k\sigma)^2 \ll 1, \quad \vec{\gamma}^2 \ll 1 \quad (2)$$

where ∇_S is a surface gradient and σ is the rms deviation of ζ , the boundary condition (1) may be transferred to the mean surface S through a Taylor series expansion. Higher order terms can be neglected by the conditions (2). We assume the fields are of the form of a mean field \vec{E} and a fluctuation field \vec{e} .

$$\vec{E} = \vec{E} + \vec{e}, \quad \vec{H} = \vec{\mathcal{H}} + \vec{h} \quad (3)$$

In this approximation, the mean field coincides with the field reflected from S , while the fluctuation field is small relative to the mean. The boundary conditions are thus rewritten

$$[\vec{n} \times (\vec{E}_1 - \vec{E}_2)]_S = 0, \quad [\vec{n} \cdot (\vec{E}_1 - \epsilon \vec{E}_2)]_S = 0, \quad \vec{\mathcal{H}}_1 = \vec{\mathcal{H}}_2|_S \quad (4)$$

$$\left. \begin{aligned} \vec{n} \times (\vec{e}_1 - \vec{e}_2) &= [\vec{\gamma} \times (\vec{E}_1 - \vec{E}_2)] - [\vec{n} \times \frac{\partial}{\partial z} (\vec{E}_1 - \vec{E}_2)] \zeta \\ \vec{n} \times (\vec{h}_1 - \vec{h}_2) &= [\vec{\gamma} \times (\vec{\mathcal{H}}_1 - \vec{\mathcal{H}}_2)] - [\vec{n} \times \frac{\partial}{\partial z} (\vec{\mathcal{H}}_1 - \vec{\mathcal{H}}_2)] \zeta \end{aligned} \right\} \quad (5)$$

Equivalent current densities \vec{j}_e and \vec{j}_m , electric and magnetic, respectively,

$$\vec{n} \times (\vec{e}_1 - \vec{e}_2) = -\frac{4\pi}{c} \vec{j}_m, \quad \vec{n} \times (\vec{h}_1 - \vec{h}_2) = \frac{4\pi}{c} \vec{j}_e \quad (6)$$

By expressing $\vec{\epsilon}_2$ in terms of $\vec{\epsilon}_1$ (through boundary conditions 4) and \vec{H} in terms of \vec{E} (through Maxwell's eqs) \vec{j}_m and \vec{j}_e are expressed in terms of \vec{E} , \vec{n} , ζ as

$$\begin{aligned}\vec{j}_e &= i k c \frac{1-\epsilon}{4\pi} [\vec{n} \times (\vec{E}_1 \times \vec{n})] \zeta \\ \vec{j}_m &= \frac{c}{4\pi} \frac{\epsilon-1}{\epsilon} [\vec{n} \times \nabla(\vec{n} \cdot \vec{E}_1) \zeta]\end{aligned}\quad (7)$$

The general field $\vec{E}_{1,2}$ at a point above (below) the dielectric is expressed as a sum of plane waves, i.e.

$$\vec{E}_{1,2}(\vec{R}) = \int_{-\infty}^{\infty} \tilde{E}_{1,2}(\vec{k}) \exp[i(\vec{k} \cdot \vec{r}) + \sqrt{k_{1,2}^2 - k^2} z] d\vec{k} \quad (8)$$

where $\vec{R} = \vec{r} + \hat{z} z$, $\vec{k} = |k| \vec{\beta}_\perp$, $\vec{\beta} = \vec{\beta}_\perp + \hat{z} \beta_z = \frac{\vec{k}}{|\vec{k}|}$

Each component plane wave must satisfy (from Maxwell's eqs)

$$\begin{aligned}\vec{h}_1 \times \vec{\beta} &= \tilde{e}_1, \quad [\vec{n} \times (\vec{h}_1 - \vec{h}_2)] = \frac{4\pi}{c} \tilde{j}_e \\ \vec{e}_1 \times \vec{\beta} &= -\vec{h}_1, \quad [\vec{n} \times (\vec{e}_1 - \vec{e}_2)] = -\frac{4\pi}{c} \tilde{j}_m \\ \vec{h}_2 \times (\vec{\beta} - (a+b)\vec{n}) &= \epsilon \tilde{e}_2, \quad [\tilde{e}_2 \times (\vec{\beta} - (a+b)\vec{n})] = -\vec{h}_2\end{aligned}\quad (9)$$

where (10)

$$\tilde{j}_{e,m}(\vec{k}) = \frac{1}{(2\pi)^2} \int \vec{j}_{e,m} e^{-i\vec{k} \cdot \vec{r}} d\vec{r}$$

and

$$a = \vec{n} \cdot \vec{\beta}, \quad b = \sqrt{\epsilon - 1 + a^2}$$

From these expressions, $\tilde{e}_1(\vec{k})$, $\tilde{e}_2(\vec{k})$ are found to be

$$\tilde{e}_1(\vec{k}) = \frac{4\pi}{c(b+a\epsilon)} \left\{ \frac{\epsilon-1}{a+b} (\vec{\beta} \cdot \tilde{j}_m)(\vec{n} \times \vec{\beta}) + [\vec{\beta} \times (\vec{n} \times \tilde{j}_e)] \right. \\ \left. + \epsilon (\vec{\beta} \times \tilde{j}_m) + [\vec{\beta} \times (\vec{\beta} \times \tilde{j}_e)] \right\} \quad (11)$$

$$\tilde{e}_2 = \tilde{e}_1 + \frac{1-\epsilon}{\epsilon} \vec{n} \cdot (\vec{n} \cdot \tilde{e}_1) + \frac{4\pi}{c} (\tilde{j}_m \times \vec{n}) + \frac{4\pi}{c\epsilon} \vec{n} (\vec{\beta} \cdot \tilde{j}_e)$$

The field \tilde{E}_1 , on the surface $z=0$ is written using the Fresnel formulas for an incident field \vec{E}^0

$$\text{where } \vec{E}^0 = \hat{p} e^{ikR_0} e^{i(\vec{k}_1 \cdot \vec{r} + k_z z)} \quad (12)$$

where R_0 is the distance from the source to the center of the scattering surface and \hat{p} is the polarization vector. Thus it can be derived

$$\vec{E}_1 = \frac{2a_0}{a_0+b_0} \vec{E}^0 - \frac{2a_0(1-\epsilon)}{b_0+\epsilon a_0} \left(\vec{n} + \frac{\vec{\alpha}}{a_0+b_0} \right) (\vec{E}^0 \cdot \vec{n}) \quad (13)$$

where

$$a_0 = -\vec{n} \cdot \vec{\alpha}$$

$$b_0 = \sqrt{\epsilon - \cos^2 \psi}$$

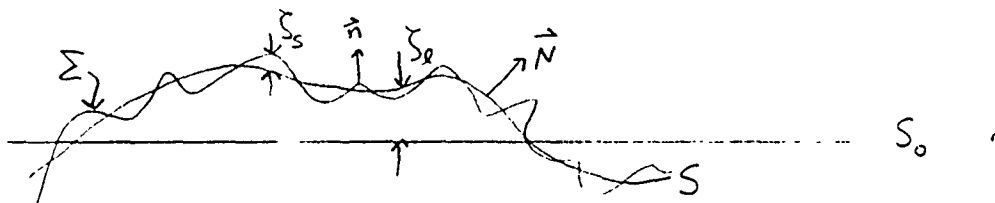


Substituting eq (13) for \vec{E}_1 into eqs (7) for \tilde{j}_e and \tilde{j}_m ; eqs (7) into eqs (10) for \tilde{j}_e and \tilde{j}_m ; and eqs (10) into eq (8), the field is found at any point in space. The integral may be evaluated by saddle-point integration in view of the fact that the observation point is located in the Fraunhofer zone. For the backscatter case where $\vec{k}_i = -\vec{k}_r$,

$$\vec{E} \cdot \hat{p} = \frac{k^2(\epsilon-1)}{\pi R_0} e^{2ikR_0} E^0 \int_A \left\{ \left(\frac{a}{a+b} \right)^2 (\hat{p} \cdot \hat{p}_0) + \frac{2(\epsilon-1)a^2b}{(b+a\epsilon)^2(a+b)} \right. \\ \left. \cdot (\vec{N} \times \hat{p})(\vec{N} \times \hat{p}_0) \right\} e^{2ik\vec{\beta} \cdot \vec{r}} \zeta(\vec{r}) d\vec{r} \quad (14)$$

For the case of a small scale only the zeroth order field \vec{E} is simply the

field reflected from a plane interface, found using Fresnel coefficients. When the SPM is applied to a composite surface, the methodology is similar, but more cumbersome. For the two scale case we have a surface, denoted by Σ , represented by normal deviations $\zeta(\vec{r}_s)$ from a smoother surface S .



The position of points on Σ is related to points on S by $\vec{r}_\Sigma = \vec{r}_s + \vec{N}\zeta(\vec{r}_s)$, where \vec{N} is the normal to the mean surface S . The normal to Σ will be denoted \vec{n} . For a medium of arbitrary permittivity ϵ below the surface and vacuum above, the boundary conditions are

$$\begin{aligned} [\vec{n} \times (\vec{E}_1 - \vec{E}_2)]_\Sigma &= 0, \quad [\vec{n} \cdot (\vec{E}_1 - \epsilon \vec{E}_2)]_\Sigma = 0 \\ (\vec{H}_1 - \vec{H}_2)|_\Sigma &= 0 \end{aligned} \quad (15)$$

Again, assuming the surface deviations from S are small, with gentle slopes such that

$$(k\sigma)^2 \ll 1, \quad (\nabla_s \zeta)^2 \ll 1 \quad (16)$$

where $\sigma^2 = \text{Var} [\zeta]$ relative to S , then an expansion of the fields in terms of powers of ζ is valid, keeping only the first two terms

$$\begin{aligned} \vec{E}(\vec{r}_\Sigma) &\approx \vec{E}(\vec{r}_s) + \zeta(\vec{r}_s)(\vec{N} \cdot \nabla)\vec{E}(\vec{r}_s), \quad \vec{n} \approx \vec{N} - \vec{\gamma} \\ \vec{H}(\vec{r}_\Sigma) &\approx \vec{H}(\vec{r}_s) + \zeta(\vec{r}_s)(\vec{N} \cdot \nabla)\vec{H}(\vec{r}_s), \quad \vec{\gamma} = \nabla_s \zeta(\vec{r}_s) \end{aligned} \quad (17)$$

The scattered field is again assumed to be of the form $\vec{E} = \vec{E}_0 + \vec{e}$, $\vec{H} = \vec{H}_0 + \vec{h}$ and the analysis proceeds in a similar fashion to the analysis for a small scale surface only. In both cases the final form for the field scattered from the small scale roughness is the same, the only difference being $k_z \zeta_0$ appears in the exponential $e^{-j\vec{k} \cdot \vec{R}}$ for the composite surface case. If the limiting case is taken, $\lim_{\zeta \rightarrow \infty}$ (perfectly conducting surface), then the field becomes

$$\vec{e} = \frac{E^0 k^2}{\pi R_0} e^{-jkR_0} \iint \frac{2(\vec{n} \cdot \hat{p})(\vec{n} \cdot \hat{p}_0) + (\vec{n} \cdot \vec{k}_i)^2 (\hat{p} \cdot \hat{p}_0)}{\sqrt{1 + (\nabla_s \zeta_0)^2} - 1} - \zeta_s e^{-j2(\vec{k}_i \cdot \vec{r} + \zeta_0 k_z)} dx dy \quad (18)$$

a simpler form of (14) to work with. The zeroth order field is found in a much more direct fashion to be [12]

$$\vec{E}_0 = \frac{-jkE^0 \delta_{pp'}}{2\pi R_0 \cos \theta} e^{-jkR_0} \iint e^{-j2(\vec{k} \cdot \vec{r} + k_z \zeta_0)} dx dy \quad (19)$$

for the perfectly conducting surface case.

DEVELOPMENT OF THE SIRV FORM

In general the SPM applied to composite surface of large and small scale variations ζ_0 and ζ_s respectively, predicts a scattered field

$$\vec{E}^s = \vec{E}_0 + \vec{e}$$

i.e. a sum of a zeroth order field and a first order fluctuation field.

Brown derives these components (for a perfectly conducting surface) as:

$$\Omega_{PP'} = \frac{E^0 k^2}{\pi r} \exp(-jkr) \iint_A \frac{2(\vec{n} \cdot \hat{p})(\vec{n} \cdot \hat{p}') + (\vec{n} \cdot \vec{k}_i)^2 (\hat{p} \cdot \hat{p}')}{\sqrt{1 + \zeta_{ex}^2 + \zeta_{ey}^2} - 1} \cdot \zeta_s \exp(-j2\vec{k}_i \cdot \vec{r}_o) dx dy$$

$$\epsilon_{opp'} = \frac{-jE^0 k \delta_{pp'}}{2\pi r \cos \theta} \exp(-jkr) \iint_A \exp(-j2\vec{k}_i \cdot \vec{r}_o) dx dy$$

where E^0 is the amplitude of incident field

\vec{k}_i specifies the propagation vector of the incident field

$\vec{r}_o = \hat{x}x + \hat{y}y + \hat{z}\zeta_z$ is the position vector from the origin to any point on the large scale surface

\vec{n} is the unit normal to the large scale surface

$k = |\vec{k}_i|$ is the wavenumber

\hat{p} , \hat{p}' denote incident and backscattered polarizations respectively

$$\zeta_{ex} = \frac{\partial \zeta_e}{\partial x}, \quad \zeta_{ey} = \frac{\partial \zeta_e}{\partial y}$$

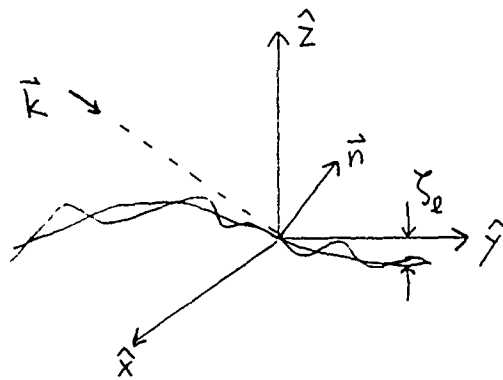
$\hat{k}_i = \frac{\vec{k}_i}{k}$ a unit vector in the direction of incident propagation

r is the distance from the transmitter-receiver to the origin

$$\delta_{pp'} = \begin{cases} 1 & \text{if } \hat{p} = \hat{p}' \\ 0 & \text{if } \hat{p} \neq \hat{p}' \end{cases}$$

A represents the beam illuminated area (radar footprint)

θ is the angle of incidence = $\angle(-\vec{k}_i, \hat{z})$



The incident field $\vec{E}_i = \hat{p} E^0 e^{-j \vec{k}_i \cdot \vec{r}}$ is assumed to be a plane wave with time dependence $j\omega t$ suppressed.

As previously mentioned the term \vec{C}_0 represents the scattered field from the large scale surface variation only. It is well known that the \vec{C}_0 term is strongest in the specular direction and falls off rapidly outside a cone centered axially on the specular wavevector. When this general result is specialized to radar backscatter the implication is that the zeroth order field has its strongest effect on the return signal at or near normal incidence $\Theta = 0^\circ$. It will in fact dominate the radar return if the radar receiver is located within the cone. It can also be shown [10, chap 5] that the spread of the cone is proportional to rms height cone width (in degrees) $\propto \sigma/\lambda$. In the limiting case as $\sigma \rightarrow 0$ the surface becomes a plane and reflection occurs in the specular direction only (Snell's Law), which is intuitively satisfying.

Now the desired mathematical form sought is a product of a Gaussian random variable and some scalar. To simplify the form of the backscattered field we assume that the slope m of the large scale variation is less than 0.2. Ulaby, Moore and Fung [13, chap 12] have calculated backscattering coefficients due to large scale roughness using the tangent plane method which shows

that large scale backscatter becomes very small for angles of incidence greater than 30° if the large scale rms slope is less than 0.2. Under these conditions, the first order field \vec{E} will dominate the radar backscatter. By grouping terms of \vec{E} which depend only on ζ_ℓ , ζ_s , or $\nabla\zeta_\ell$, the scattered field \vec{E}^s has the form (subject to the aforementioned limitation)

$$\vec{E}^s \approx \vec{E} = C \iint_A \Gamma_{pp'}(\nabla\zeta_\ell) \zeta_s e^{-j2(\vec{k} \cdot \vec{r}_{xy} + k_z \zeta_\ell)} dx dy$$

where $C = \frac{E^0 k^2}{\pi r} e^{-jk r}$

$$\Gamma_{pp'}(\nabla\zeta_\ell) = [2(\vec{n} \cdot \hat{p})(\vec{n} \cdot \hat{p}') + (\vec{n} \cdot \vec{k}_i)^2 (\hat{p} \cdot \hat{p}')] \sqrt{1 + \zeta_{\ell x}^2 + \zeta_{\ell y}^2}$$

The basic area of integration A for a single look of the radar (sometimes called the radar footprint) can be broken up into a summation over smaller subareas or patches. Since ζ_ℓ and $\nabla\zeta_\ell$ vary much more "slowly" over the footprint, a logical way to break up the integral is by patches over which ζ_ℓ and $\nabla\zeta_\ell$ remain fairly constant. Doing so yields

$$\vec{E}^s = C \sum_p \Gamma[\zeta_\ell(p)] e^{-j2k_z \zeta_\ell(p)} \iint_{\Delta p} \zeta_s(x,y) e^{-j2(\vec{k} \cdot \vec{r}_{xy})} dx dy$$

or more compactly

$$\vec{E}^s = \sum Z(p) \iint_{\Delta p} \zeta_s(x,y) e^{-j2(\vec{k} \cdot \vec{r}_{xy})} dx dy$$

where

$$Z(p) = C \Gamma[\zeta_\ell(p)]$$

\vec{E}^s is now of the form

$$\vec{E}^s = Z_1 \iint_{\Delta 1} [\quad] dx dy + Z_2 \iint_{\Delta 2} [\quad] dx dy + \dots + Z_n \iint_{\Delta n} [\quad] dx dy$$

Now consider a general term of the series

$$Z_P \iint_{\Delta_P} \zeta_s(x, y) e^{-j2(\vec{k} \cdot \vec{r}_{xy})} dx dy$$

If the size of $\Delta_P \gg \ell^2$ where ℓ is the correlation length of ζ_s , then the integral

$$\iint_{\Delta_P} \zeta_s e^{-j\vec{k} \cdot \vec{r}_{xy}} dx dy$$

can also be expressed as a sum over smaller areas

$$\iint_{\Delta_P} \zeta_s e^{-j2\vec{k} \cdot \vec{r}_{xy}} dx dy = \sum_m \iint_{\Delta_m} \zeta_s e^{-j2\vec{k} \cdot \vec{r}_{xy}} dx dy = \sum_m X_m$$

Each X_m is a random variable depending on ζ_s . By choosing the subpatch area Δ_m as a square of dimensions $\ell \times \ell$, then each subpatch will be independent of the others. Neighboring subpatches would be somewhat correlated, but here we will assume that the effect is small and neglect it. The mean of X_m

$$\langle X_m \rangle = \iint_{\Delta_m} \langle \zeta_s \rangle e^{-j2\vec{k} \cdot \vec{r}_{xy}} dx dy = 0$$

since by definition $\langle \zeta_s \rangle = 0$.

The variance of X_m

$$\langle X_m^2 \rangle = \iiint \iiint \langle \zeta_s(x, y) e^{-j2\vec{k} \cdot \vec{r}} [\zeta_s(u, v) e^{-j2\vec{k} \cdot \vec{w}}]^* \rangle dx dy du dv$$

$$= \iiint R_s(x-u, y-v) e^{-j2[k_x(x-u) + k_y(y-v)]} dx dy du dv$$

by stationarity of \tilde{E}_s .

This fourfold integral depends only on the differences $x-u$, $y-v$ and will be the same for every subpatch. If the number of subpatches is large enough, then by the central limit theorem, $\sum_m \chi_m$ is a zero mean Gaussian random variable G , the same for all patches. Or,

$$\tilde{E}^s = G \sum_p Z_p = G \cdot S$$

$$\text{where } S = \sum_p Z_p.$$

This is precisely the form of an SIRV. As the radar beam moves in time, either by scanning or as it is attached to an aircraft, the returns will be samples of a process, an SIRP.

SUMMARY

A radar clutter model derived from electromagnetic principles has been shown to have the proper form of a spherically invariant random process. Characterization of radar clutter as an SIRP from theory may provide an important link to predicting the specific clutter statistics based on terrain topology and material parameters (μ, ϵ) that make up the surface. It should be noted, however, that a more general formulation of a radar clutter model should include volume scattering (by vegetation) as well as surface scattering, and also provide for near normal incidence and near grazing incidence.

REFERENCES

- [1] Guinard, N.W., Ransone, J.T., Laing, M.B., and Hearton, L.E.: 'NRL terrain clutter study, phase I', Naval Research Laboratory, NRL report 6487, May 1967
- [2] Fay, F.A., Clarke, J., and Peter, R.S.: 'Weibull distribution applies sea clutter'. IEE Con. Publ. 155 (Radar 77), 1977 pp.101-102
- [3] Trunk, G.V., and George, S.F.: 'Detection in non-Gaussian sea clutter'. IEEE Trans., 1970, AES-8, pp. 620-628
- [4] Jakeman, E., and Pusey, P.N. : 'A model for non-Rayleigh sea echo', 1976, AP-24, pp. 806-814
- [5] Jakeman, E.: 'On the statistics of K-distributed noise'. J. Phys. A, 1980, 13, pp. 31-48
- [6] Conte, E., Longo, M.: 'Characterisation of radar clutter as spherically invariant random process'. IEE Proc., Vol. 134, Pt. F, No. 2, April 1987 pp. 191-197
- [7] Yao, K.: 'A representation theorem and its applications to spherically-invariant random processes', IEEE Trans., 1973, IT-19, pp. 600-608
- [8] Goldman, J.: 'Detection in the presence of spherically symmetric random processes', 1976, IT-22, pp. 52-58
- [9] Brehm, H.: 'Description of spherically invariant random process means of G-function', Springer Lecture Notes, 1982, 969, pp. 39-72
- [10] Beckmann, P., Spizzichino, A.: 'The scattering of electromagnetic waves from rough surfaces', New York, Pergamon Press, 1963
- [11] Brekhovskikh, L'v.: 'The diffraction of waves by a rough surface, part I, Zh. Eksper. i. Teor. Fiz. 23, 1952 pp. 275-289
- [12] Brown, G.S.: 'Backscattering from a Gaussian-distributed perfectly conducting rough surface', IEEE Trans., AP-26, 3, 1978 pp. 472-482
- [13] Ulaby, F.T., Moore, R.K., Fung, A.K.: 'Microwave remote sensing - active and passive': Vol II, Addison-Wesley, 1982
- [14] Rice, S.O.: 'Reflection of electromagnetic waves from slightly rough surfaces', Comm. Pure Appl. Math. 4, pp 351-378, 1951
- [15] Burrows, M.L., 'A reformulated boundary perturbation theory an electromagnetic and its application to a sphere', Can. J. Phys., vol. 45, pp. 1729-1743, May 1967
- [16] Burrows, M.L., 'On the composite model for rough surface scattering', IEEE Trans. AP-21, pp. 241-243, Mar 1973

- [17] Beckmann, P.: 'Scattering by composite rough surfaces', Proc. IEEE, Vol 53, pp. 1012-1015, Aug 1965
- [18] Fung, A.K., Chan, Hsiao-Lien,: 'Backscattering of waves by composite rough surfaces', IEEE Trans. AP-17, 5, Sept 1969
- [19] Bass, F.G., Fuks, I.M., La;ukov, A.I., Ostrovsky, I.E., Rosenberg, A.D., 'Very high frequency radiowave scattering by a disturbed sea surface, part II', IEEE Trans. AP-16, 5, pp. 560-568, Sept 1968
- [20] Bass, F.G., Fuks, I.M.: 'Wave scattering from statistically rough surfaces', Inst. Radiophy. Electronics, Kharkov, USSR, translated:Vešecky, C.B., Vesecky, J.F. pp. 418-442 Pergamon Press 1979
- [21] Valenzuela, G.R.: 'Scattering of electromagnetic waves from a tilted slightly rough surface', Radio Sci., Vol. 3, no. 11, pp. 1057-1066, Nov 1968

Final Report
Summer Graduate Student Research

By
Kuang Z. Cheng

July 1, 1991 to September 6, 1991

Focal Point : *Harold Rott*

Introduction:

We have set up three systems (Fig. 1), namely, Laser & Optics, Atomic Beam, Microwave systems to do the following diagnostics and experiments.

----- Optical Pumping

----- Direct Excitation of Microwave-Spin Dressed States
using a Laser-Excited Resonance Raman Interaction
(Ref. 1)

----- Raman Induced Microwave Echo in an Atomic Beam

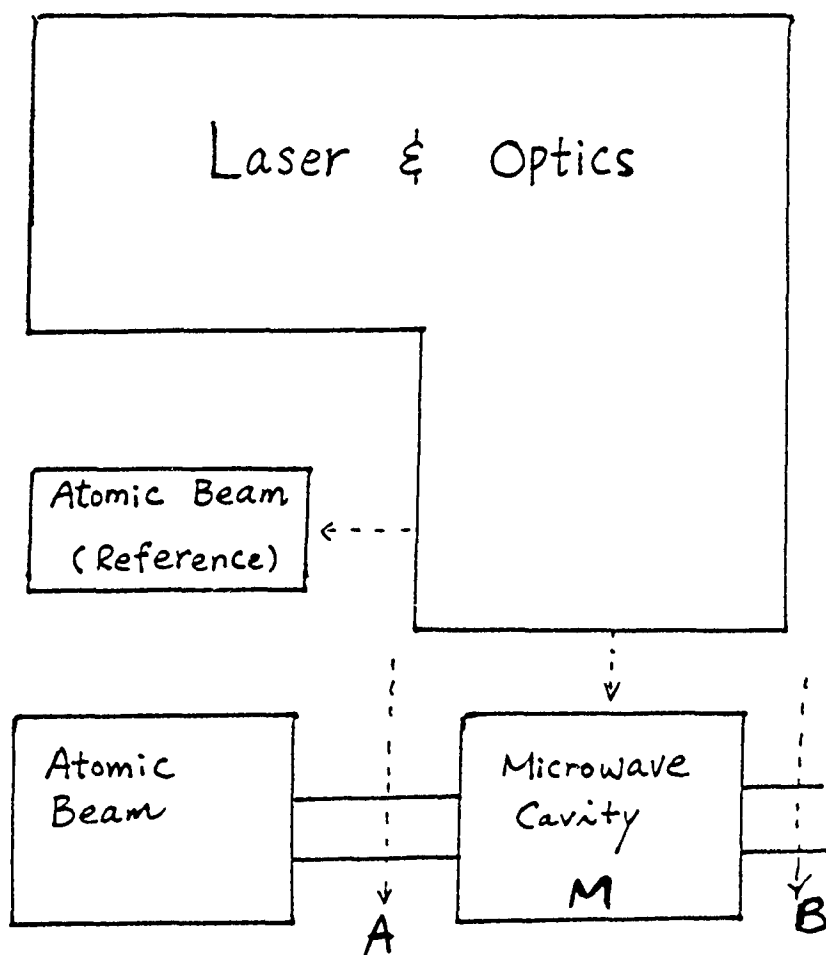


Fig. 1

For the Optical Pumping experiment, we pre-pump the Sodium beam by laser going through the window of section A. Since the light irradiation is equivalent to a time-dependent stochastic process acting on the atoms, the excited Na atoms decay to level 2 with finite probability. Therefore, the fluorescence after laser going through window of section B will be reduced.

For the experiment of Excitation of Microwave-spin Dressed States, we pump the atoms in zone A into the trapped state, i.e. Raman excitation. These excited atoms interact with a microwave field in zone M. We measure the degree of this interaction in zone B by detecting the loss of the trapped state population, via the fluorescence detecting photodiode. We will also find out the phase effect between microwave field and optical difference frequency.

For the typical photon echo experiment, the first (Object) and second (Reference) pulse interfere to form a grating which is stored as spectrally burned holes in the ground state population. At a later time, the Read pulse diffract off this grating resulting in a reproduction of the Object pulse after a time equaling the time separation between the first two pulses.

In the Raman Induced Spin Echo experiment, A Raman (two optical frequencies, resonant with the two optical transitions) pulse carries the Object image, and a microwave pulse is the Reference. The Raman pulse excites the magnetic spin dipoles which interfere with the microwave pulse to form gratings. At a later time, a third microwave pulse would temporally diffract off this grating, generating a microwave (spin) echo of the same shape as the Object Raman pulse.

In the Raman process, the microwave dipoles are generated through optical pumping. Thus, the amount of time needed to excite these microwave dipoles is limited by the decay time of the atoms. So, the system uses pulses that are several orders of magnitude longer than the pulse necessary in conventional photon echo, and yet has similar data number.

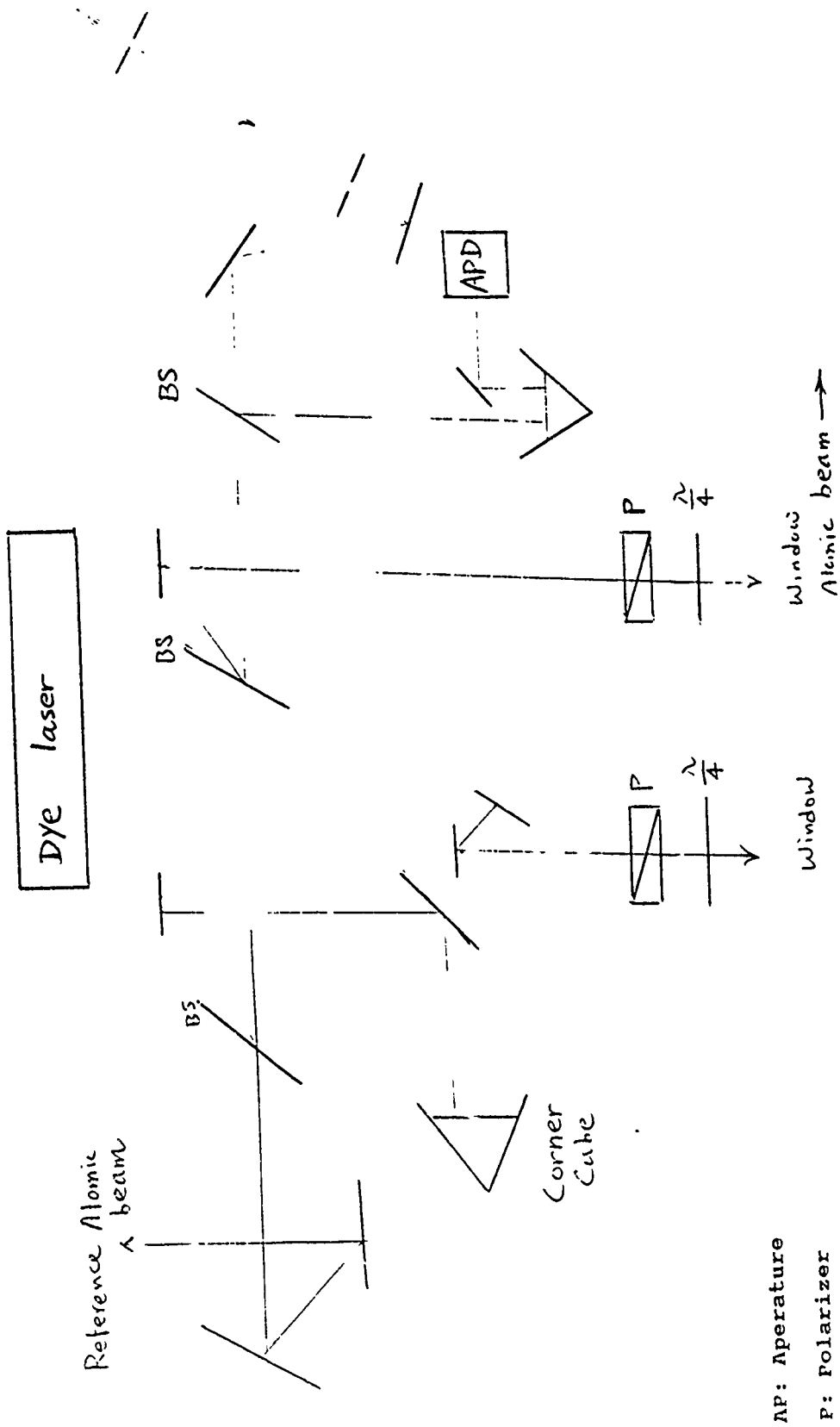
The necessity of inhomogeneous broadening on the atomic beam in echo experiment is achieved by using the inhomogeneous magnetic field.

(1) Laser & Optical System

We set up the Dye Laser (Spectra-Physics 380, 540--650nm, 500 mW) which is pumped by Argon Ion Laser (Spectra-Physics 2030, 458-- 514 nm, 20W).

The beam of cw Dye laser can be locked or scanned by setup of Fig.2.

Fig.3 is the setup of optical system which guides the Dye laser beam through the windows of atomic beam.



AP: Aperature

P: Polarizer

BS: Beam Splitter

FP: Fabry-Perot Etalon

PD: Photodiode

TS: Telescope

APD: Avalanche Photodiode

 $\lambda/4$: Retardation Plate

Fig. 3

(2) Atomic Beam

We have used sodium which is the three-level atomic Λ configuration as the source of beam. The sodium oven (Fig.4) is wrapped by heating wire and insulation tape.

For the reference beam, the end section of oven is operated at about 145°C and the middle section is 485°C to generate a thermal atomic beam of 2 mm diameter.

The ~~supersonic~~ ^{high-density} beam is generated by the construction of Fig.5. The temperature of oven are 370°C for front section and 480°C for rear section.

The whole system is under mechanical and diffusion pumping. The liquid-nitrogen-fill trap is built on the cavity to reduce the pressure. The leak check is accomplished in advance.

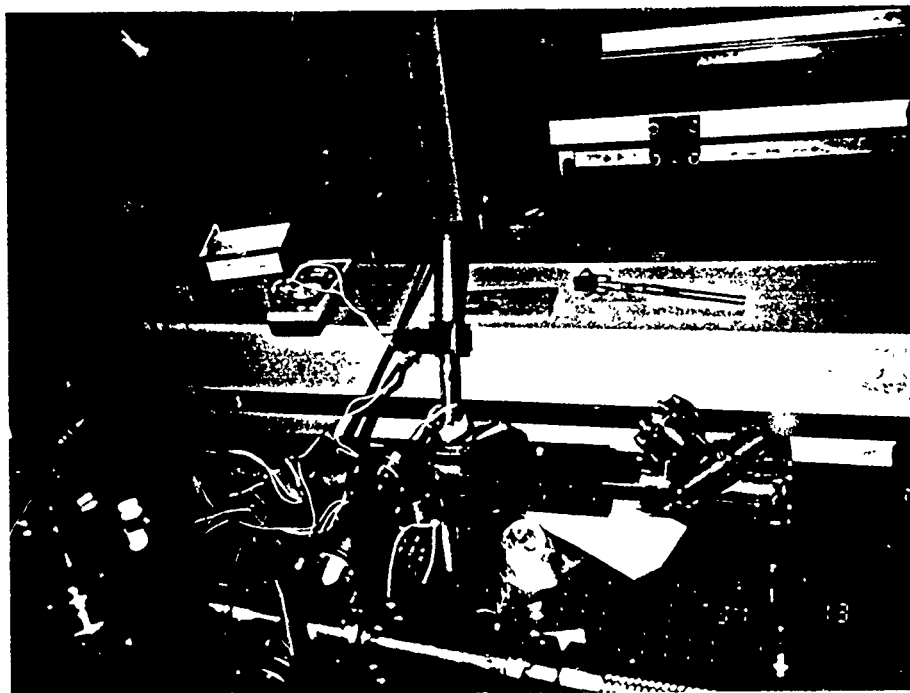


Fig. 4

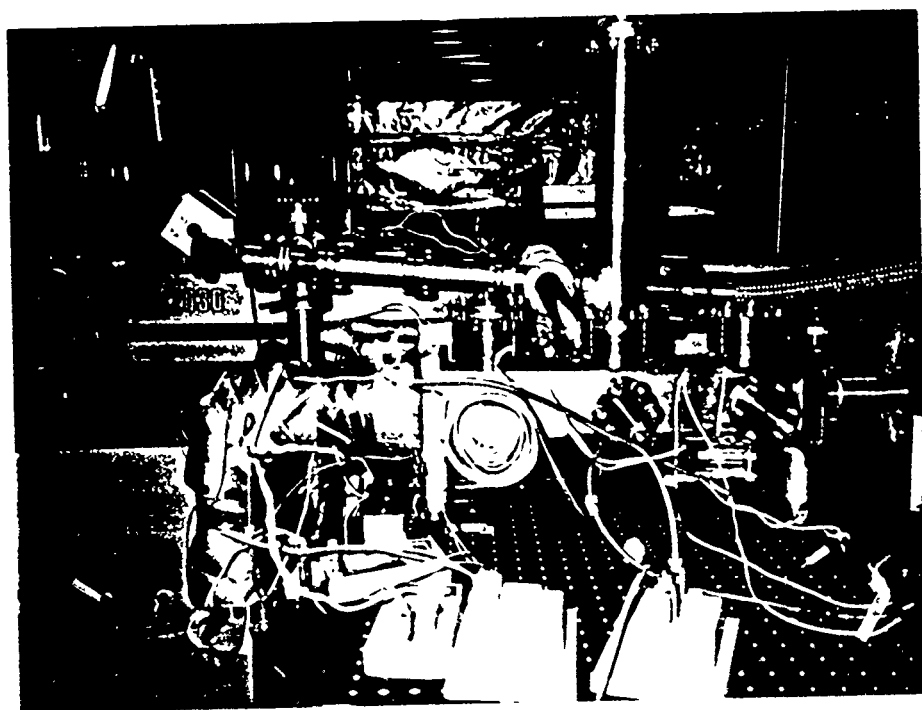


Fig. 5

(3) Microwave system

The cavity, 11.2cm X 5.8cm X 29.8cm, is designed to have TE₁₀₂ mode and resonance frequency of 1772 Mhz which is the 1 \rightarrow 3 transition frequency of sodium beam. (Fig.6).

The electric E field and magnetic B field of the cavity are drawn in Fig.7 according to the following equations.

(Ref.4)

$$\vec{E} = \hat{y} E_{mp} \sin \frac{m\pi}{a} x \sin \frac{p\pi}{d} z$$

$$\vec{H} = \hat{x} E_{mp} \sin \frac{m\pi}{a} x \cos \frac{p\pi}{d} z - \hat{z} E_{mp} \frac{m\pi}{j\omega\mu a} \cos \frac{m\pi}{a} x \sin \frac{p\pi}{d} z$$

In order to have inhomogeneous B field in the cavity, we made the loop coil to generate gradient B field of 6 Gauss per centimeter. (Fig.8) The number of coil turns and current are determined by the equation of (Ref.4)

$$\Delta B = \frac{\mu N I}{2} \frac{a^2}{(a^2 + z_o^2)^{\frac{3}{2}}} \frac{6z_o}{a^2 + z_o^2}$$

where N is the number of turns, a is the radius of the loop,

z_o is the half of separation.

We also wired 21 turns of big loop across the whole cavity system to provide optional homogeneous B field. (Fig.9)

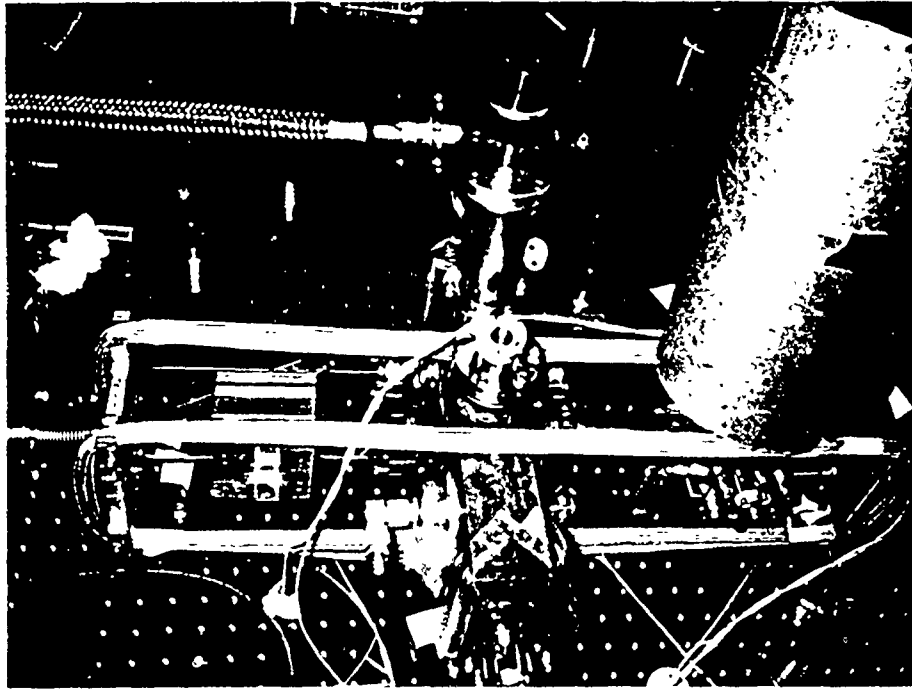
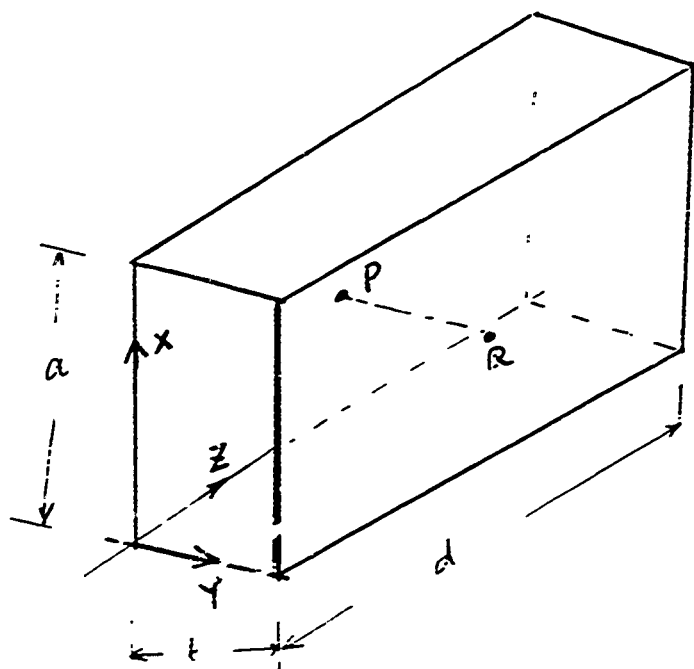


Fig. 6 .



The field from
 $P(\frac{a}{2}, 0, \frac{d}{2})$ to
 $Q(\frac{a}{2}, b, \frac{d}{2})$

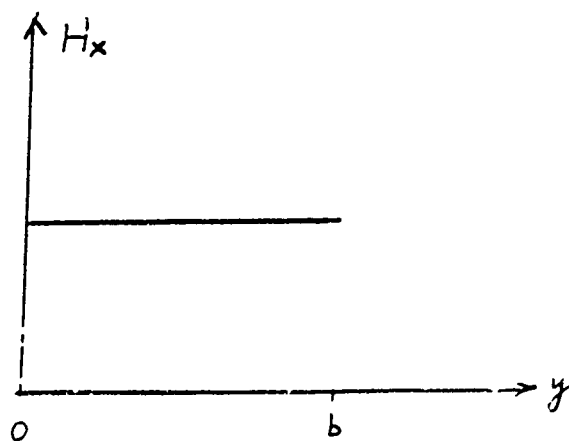
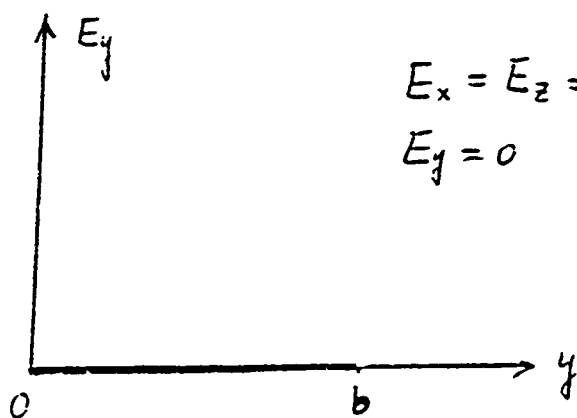


Fig 7

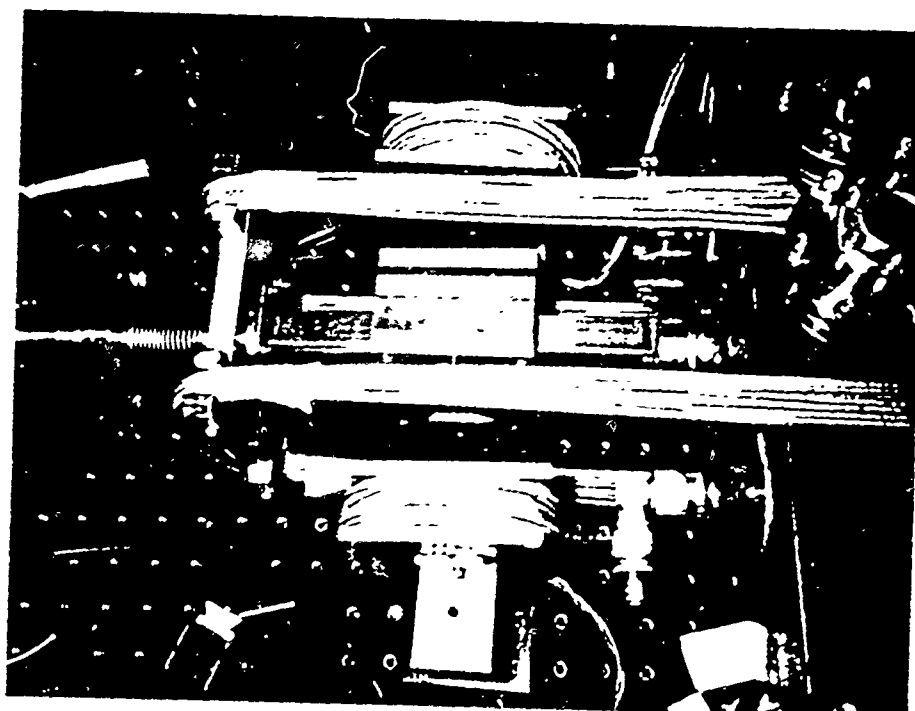
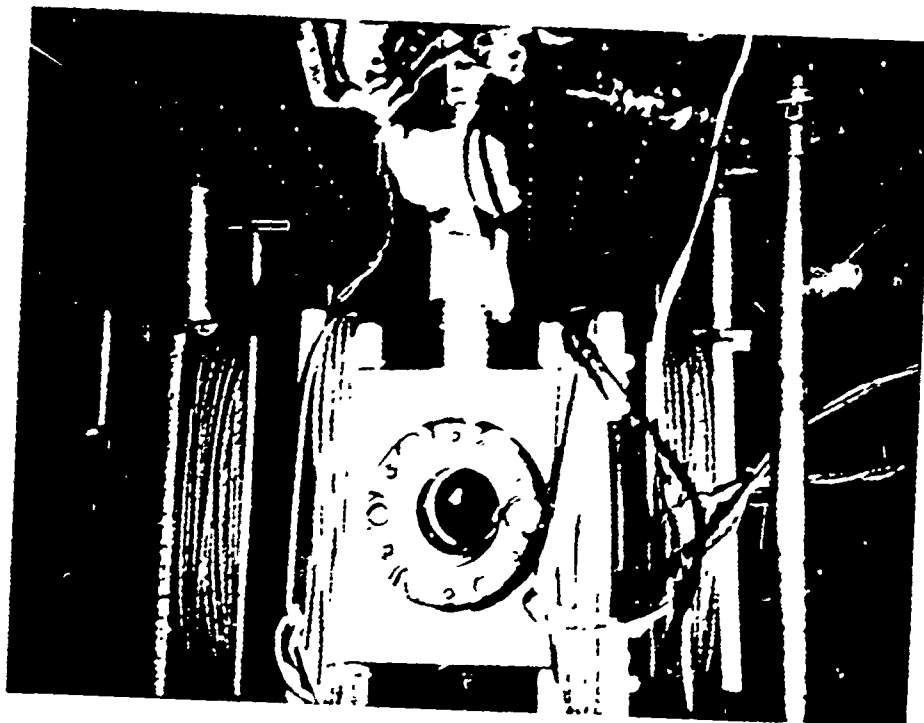


Fig. 8

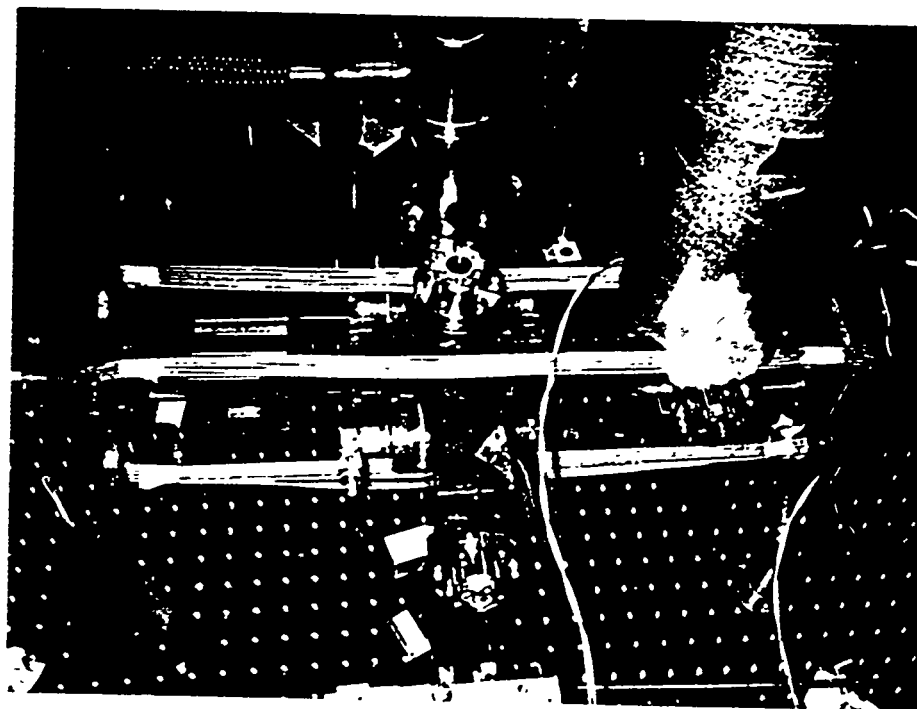
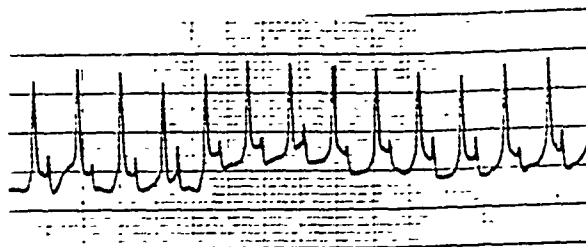
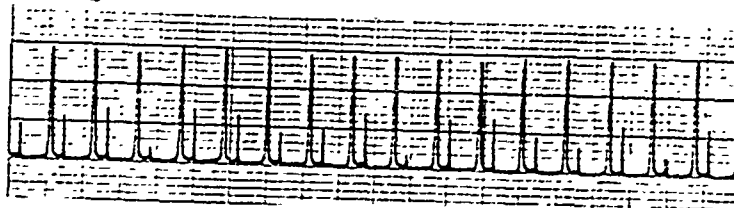


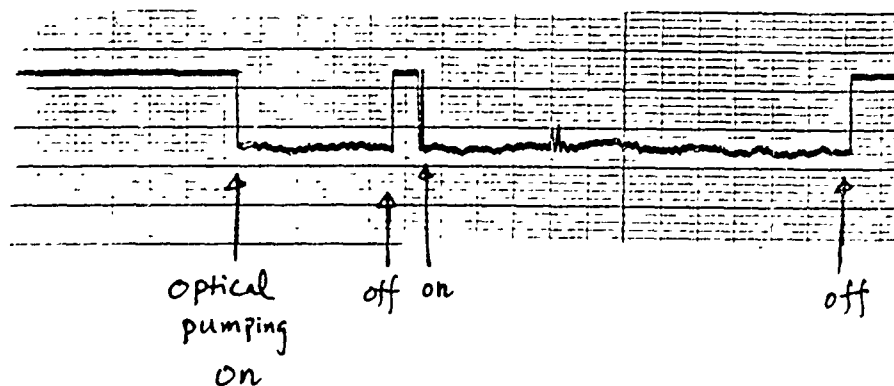
Fig. 9

Conclusion:

The atomic beam of 2 mm diameter has been generated by 25 microampere laser beam. The chart record is following.



Also the optical pumping is recorded as following.



This is a valuable experience for me. Especial thanks to Dr. Phil Hemmer, Salin Shahriar, and John Kierstead for their kind teaching and guidance.

Ref. 1 M.S. Shahriar and P.R. Hemmer, Phys. Rev. Lett.
65, 1865 (1990)

Ref. 2 D. Suter, M. Rosatzin, and J. Mlynek, Phys. Rev.
Lett. 67, 34 (1991)

Ref. 4 L.C. Shen and J.A. Kong, Applied Electromagnetism
_____ (PWS Publishers, Massachusetts, 1987)

1991 USAF-RDL SUMMER RESEARCH PROGRAM

**MILLIMETER-WAVE NOISE
MODELING INVESTIGATION**

**Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Administered by the
Research and Development Laboratories**

FINAL REPORT

| | |
|-------------------------------|--|
| Faculty participant: | Lawrence P. Dunleavy, Asst. Professor |
| Graduate student participant: | Steven M. Lardizabal |
| University Affiliation: | Department of Electrical Engineering University of South Florida Tampa, FL 33620 |
| USAF Researcher: | Richard Webster Rome Laboratories, RL/EEAC Hanscom AFB, MA 01731 |
| Date: | July 19, 1991 |
| Contract No.: | F49620-90-C-0076 |

MILLIMETER-WAVE NOISE MODELING INVESTIGATION

ACKNOWLEDGEMENTS

This work was sponsored under the United States Air Force AFOSR Summer Research Program, (Contract #F49620-90-C-0076) administered by Research and Development Laboratories.

The encouragement and assistance of Mr. R. Webster and Dr. P. Carr, and the endorsement and support of Dr. J. Schindler and Dr. R. Mailloux, and Lt. Col. R. Momberg are gratefully acknowledged.

We would also like to acknowledge the helpful insights gained through conversations with the following noise modeling experts: Dr. R. Pucel of Raytheon Company, Dr. M. Pospieszalski of the National Radio Astronomy Observatory, Dr. M. Gupta of Hughes Aircraft Company, and Mr. V. Adamian of Automatic Testing and Networking (ATN) Company. ATN also helped provide the noise measurements reported here.

MILLIMETER-WAVE NOISE MODELING INVESTIGATION

ABSTRACT

A new method is described that provides valid mm-wave noise models for field effect transistors, including MESFETs and HEMTs. This method avoids the complications of a variable impedance tuner, and requires only the knowledge of a small signal equivalent circuit, and noise figures measured across a range of frequencies for a single known source impedance. Noise parameters derived from this method are shown to agree well with those obtained from tuner based measurements. A review of previously reported noise modeling techniques, summarized here, reveals either their common dependence on tuner based measurements or the use of approximations that are not valid at mm-wave frequencies.

MILLIMETER-WAVE NOISE MODELING INVESTIGATION

I. INTRODUCTION

It is well known that the use of a low noise amplifier (LNA) at the front end of a mm-wave radar or communications receiver can improve system sensitivity and dynamic range as compared to the use of down conversion with a mixer immediately following the antenna. LNAs are now possible at frequencies as high as W-band (75-100GHz), using state-of-the-art High Electron Mobility Transistors (HEMTs). However, accurate transistor noise characterization is critical to achieving good millimeter-wave (mm-wave) LNA performance, while avoiding the widely used process of hand tuning. Such hand tuning for noise performance, costly and time consuming for hybrid amplifiers, is simply not practical for millimeter-wave monolithic microwave integrated circuits (MMICs).

As illustrated in Figure 1, FET noise characterization can take the form of measured noise parameters ("Black Box Methods") or a noise model ("FET Specific Methods"), from which noise parameters can be calculated. A knowledge of four noise parameters allows for a calculation of the device's noise figure for a given complex input terminating admittance, according to [7]:

$$F = F_{\min} + \frac{R_N}{G_s} |Y_s - Y_{\text{opt}}|^2 \quad (1)$$

where F is the noise figure of the two-port device, Y_s is the input or source admittance presented to the device, and G_{opt} , B_{opt} , F_{\min} , and R_N , are the four noise parameters defined as follows:

$Y_{opt} = G_{opt} + jB_{opt}$ is the optimum input admittance¹

F_{min} is the minimum noise figure achieved for $Y_s = Y_{opt}$

R_N is called the equivalent noise resistance, and describes how rapidly the noise figure degrades as the source admittance moves away from Y_{opt} .

Noise parameter measurements are relatively routine at microwave frequencies, using conventional techniques [1-4]. However, progress has been slow towards the implementation of mm-wave noise parameter measurements. This is mainly due to the difficulty with the fabrication and characterization of variable impedance tuners at mm-wave frequencies as discussed in Section II.

The present research has explored FET specific noise model methods (Figure 1) with the goal of providing an alternative or complementary mm-wave FET noise characterization method to those relying on tuner based measurements. Accordingly, an overview of FET noise modeling methods is given in Section III. Also described (Section IV) is a new method developed for deriving valid mm-wave FET noise models through the use of an equivalent circuit and a set of noise figure measurements over frequency. Figure 2 shows the model topology used.

II. OBJECTIVES AND ACCOMPLISHMENTS

At the start of the summer's work, a short "Objectives and Summary Report" containing the following information was submitted to our Air Force sponsors. This provided the template for the Summer's work.

PROBLEM: Conventional microwave noise modeling methods utilize a variable impedance (or admittance) tuner, which presents a sequence of "known" source impedances Z_S (admittances Y_S) to a device's input terminals. The tuner must be able to present a range of impedances that represent a good coverage of the Smith Chart, including near the edges of the chart. Hence, the tuner must have very low loss. The ability to fabricate, and characterize, such a tuner becomes increasingly more difficult as frequency is increased. Research and commercial developments are being pursued in the industry and at the Air Force² implementing mm-wave tuners for noise parameter measurements. Because of the inherent difficulties with this approach, however, it is essential that alternatives be explored to the reliance on a variable impedance tuner for mm-wave FET noise characterization.

SCOPE OF RESEARCH: This research consisted of a ten week effort involving the work of two researchers. The work involved a paper study, analytical research, and numerical experiments. The numerical experiments combine analytical research with existing measurement data in an attempt to explore the validity of a new noise modeling approach (see Section IV). Some of the numerical experiments were carried out by modifying existing computer programs generated by a previous Summer Research Associate. A new computer program was also generated as part of this work. Because of the time constraints, no attempt was made to perform measurements for use with this study, instead existing data was made available to us from the industry.

SPECIFIC OBJECTIVES VS. ACCOMPLISHMENTS: The work described above was carried out as planned and all of the objectives of the

work were satisfied. This degree of accomplishment was aided by extensive interaction with, Mr. Rick Webster. This included short weekly reports and weekly meetings.

The research was divided into several tasks and accomplished as follows:

1) OBJECTIVE - Review and assess the previous MESFET noise modeling work of Mr. Bill Patience [5], as performed under the 1989 AF Summer Research Program.

ACCOMPLISHMENT - After verifying the theoretical derivations of this work, an assessment was made with regard to the theoretical assumptions made, including that of frequency independence for noise source coefficients (P,R, and C) used with the model of Figure 2. This assessment was made from theoretical considerations, and numerical case studies derived from measurement data (data which was not available to Mr. Patience at the time of his study). See Section IV³.

2) OBJECTIVE - Explore improvements and alternatives to the use of tuner based measurements for noise characterization.

ACCOMPLISHMENT - A paper study was conducted to formulate an overview of various noise modeling methods documented in the literature (Section III). This study proposes a modification to the work of Patience as the most promising mm-wave noise modeling method, that does not require tuner based measurements. A related new computer program was generated.

3) OBJECTIVE - Prepare a final report in accordance with AFOSR Summer Research program guidelines, and a "mini-grant" proposal for a follow-on research project.

ACCOMPLISHMENT - The present document satisfies the final report requirements, and a related oral presentation was given at Hanscom AFB July 12, 1991. Because this work is of interest to other Air Force Laboratories, presentations of the work are scheduled to be given at Griffiss AFB July 22, 1991, and at Wright Patterson AFB August 15, 1991. The completion of the mini-grant proposal is forthcoming.

4) OBJECTIVE - Prepare a draft (or summary) version of a research paper aimed at publication in a trade journal (e.g. *IEEE Trans. on Microwave Theory and Tech.*), a trade magazine (e.g. *Microwave Journal*), or a conference digest (e.g. *IEEE MTT-S Symposium Digest*).

ACCOMPLISHMENT - An abstract for a paper entitled "A New Millimeter-wave Noise Modeling Method for Field Effect Transistors," was submitted to the 1991 IEEE International Electron Devices Meeting (IEDM). If accepted, an expanded four page paper will be required for the conference digest. It is also expected that a second paper will be generated that gives a comprehensive review of FET noise modeling methods.

III. A BRIEF OVERVIEW OF FET NOISE MODELING METHODS

A viable approach for mm-wave noise characterization of FETs is to use knowledge of an equivalent circuit for the FET to reduce the required noise measurement complexity. Several such FET specific methods have been

proposed in the literature (Figure 1). A brief summary of these methods is given here.

The reported FET noise modeling methods [7-30] use various approximations, but follow from a common theoretical formulation. This theoretical formulation is based upon the fundamental linear two port noise theory developed by Rothe and Dahlke [6] and further developed by an IRE Subcommittee on Noise [7]. This two-port noise theory applies to both Black Box noise parameter measurements and FET Specific Noise Models as indicated in Figure 1.

EARLY FET NOISE THEORY: The early groundwork for FET noise modeling was laid by van der Ziel [8-10]. van der Ziel [13], identified the main source of noise in FETs as thermal noise in the channel. He derived useful analytical expressions for the intrinsic gate and drain current sources (i_{ng} and i_{nd}) and the correlation coefficient (C) between them. Under his approximations, C was found to be imaginary. Klaassen, in a slightly more general formulation [11], found C to be complex in general with a small real part, which he attributed to high frequency gate-channel coupling (neglected by van der Ziel). Nonetheless van der Ziel's work, complete with imaginary C , has provided the format for nearly every noise modeling solution to follow.

IMPORTANT NOISE THEORY EXTENSIONS: Important extensions to this early work were made by Baechtold [12-13]. Baechtold advanced a complete noise modeling procedure, that relies on frequency independent noise coefficients P , R , and C that follow directly from van der Ziel's relations [8-

9]. The intrinsic noise sources are expressed in terms of P,R, and C through the following relations:

$$\begin{aligned} \overline{i_{nd}^2} &= 4kT\Delta f g_m P & \overline{i_{ng}^2} &= \frac{4kT\Delta f C_{gs}^2 \omega^2}{g_m} R & jC &= \frac{\overline{i_g i_d}}{\sqrt{\overline{i_{nd}^2} \overline{i_{ng}^2}}} \end{aligned} \quad (2)$$

This same noise source description is used in the present research (Figure 2). In Baechtold's method, noise parameters are calculated using a small signal model, fitted to measured S-parameters combined with analytically determined noise coefficients P,R and C.

Another very important contribution was made by Pucel et. al. [14]. This work, covers all aspects of MESFET operation, including DC, small signal, and noise characteristics. Pucel et. al. include a review is given of prior work in each of these areas and point out that the previous FET noise theory (including that of van der Ziel and Baechtold) fails to adequately account for velocity saturation effects. The authors proceed to use a two section velocity field model, valid in the saturation region to derive expressions for P,R, and C and for the four noise parameters. This method, like that of Baechtold's, is not directly applicable for accurate millimeter-wave noise characterization due to the omission of parasitic elements C_{dg} and R_{ds} (see Figure 2) from the analysis. In addition, this model shares the limitations of other purely analytic/numerical models discussed below.

The remaining FET noise model work (Figure 1) has been placed into three categories: 1) Analytical/Numerical Noise Models, 2) Semi-empirical Noise Models, and 3) Physically Based Theoretical/Experimental Models.

ANALYTICAL /NUMERICAL NOISE MODELS: The first category (Figure 1) includes the work of van der Ziel, Baechtold, and Pucel et. al., as well as the subsequent work of Cappy et. al. [15-16], Heinrich [17-18], and Brookes [19]. These methods apply basic physics to calculate noise properties, using as inputs the geometry and DC operating conditions of the FET. Such models are very useful for establishing physically meaningful model topologies and studying trends in electrical behavior with changes in geometry and operating conditions. However, analytic/numerical methods, by themselves, do not produce the most accurate FET models for circuit design. This is due to the many non-ideal conditions existing in practical FETs including, for example, variations in the geometrical dimensions and doping levels. This information is also seldom available to the circuit designer. Hence, a practically useful FET noise model requires that at least some of the noise model parameters be determined through the use of measured data.

SEMI-EMPIRICAL NOISE MODELS: The second category of noise models (Figure 2) uses measured data to determine empirical noise fitting factors. These fitting factors are used along with a simplified equivalent circuit to estimate noise parameters. Most notable of these are the models proposed by Fukui [20-22], and Podell[23]. These models are useful in that they provide convenient closed form relations between dominant geometrical or equivalent circuit parameters and noise performance, yet because of the many

approximations made they are not well suited for mm-wave FET noise characterization.

PHYSICALLY BASED THEORETICAL/EXPERIMENTAL MODELS:

The third category (Figure 1) contains the most promising methods for mm-wave applications. This work includes the methods of Gupta et. al. [24-25], Pospieszalski et. al. [26-27], Robertson et. al. [28-30], and Riddle [31]. All of these methods assume (or equivalently assume) that the coefficients P , R , and C are independent of frequency and rely on the use of measured S-parameters to derive a small signal equivalent circuit. The methods differ in the characterization of the noise sources within the model, in the simplifying assumptions and the exact model topologies.

Gupta et. al. [24-25] use a simplified model that only requires the output noise power spectral density that is determined through an output noise power measurement made at a single low frequency (e.g. 1-2GHz). Pospieszalski et. al. [26] use a more general model that requires a set of noise parameters at a single frequency to characterize equivalent temperatures T_g and T_d used to represent the gate and drain noise sources respectively. Robertson et. al. [28-30] also use a fairly general model, similar to that of this research (Figure 2). They determine the noise coefficients P , R , and C by fitting an equation for F_{min} , given in terms of P , R , and C to a set of measured F_{min} values determined for a number of frequencies (≥ 4). Most recently, Riddle [31] presented a means to directly extract P , R , and C values for a model similar to Figure 1 by using sequential matrix manipulations of the measured single frequency noise parameters and S-parameters.

Of these "Physically Based Theoretical/Experimental Noise Models," methods, the most extensively implemented and verified is that of Pospieszalski's [27], and this method has been applied to develop noise models for use at cryogenic temperatures [26]. Nonetheless, while each of the methods discussed have merit, due to various approximations they are generally not valid at millimeter-wave frequencies, or if valid require tuner based measurements in order to characterize the internal noise sources of the model.

IV. A NEW MILLIMETER-WAVE NOISE MODELING METHOD

The method advanced as part of the present research, overcomes the difficulties of the above methods. The new method makes use of noise figure data measured over a range of frequencies along with the equivalent circuit parameters to determine three frequency independent noise coefficients (P,R, and C). This technique avoids unnecessary simplifications to the equivalent circuit model and remains valid at millimeter-wave frequencies.

METHOD OF PATIENCE: The new method builds on the method described by W. Patience [5], which is illustrated in the algorithm of Figure 3. The method requires a small signal model, derived from S-parameter measurements, and noise figure data taken over a range of frequencies for a single known source reflection coefficient condition Γ_s , also measured over frequency. Central to the algorithm is an equation for the noise figure expressible in the following form:

$$F = M_1 + M_2 R + M_3 P + M_4 \sqrt{RP} \quad (3)$$

where M_i , $i = 1,2,3, 4$, are functions of the known equivalent circuit parameters, and measured values of Γ_s , corresponding to each measured value of noise

figure supplied. The unknowns, P , R , and C , are determined using a least squares fitting algorithm [2], equation (3), and the measured data. Once they have been determined, a synthesized tuner based "measurement" procedure computes the noise parameters at any desired frequency. The noise figure and reflection coefficient data required by this algorithm can readily be obtained from on-wafer or fixtured measurements even at millimeter-wave frequencies. An example of an applicable mm-wave measurement system and procedure is described by Dunleavy [32].

EXAMINATION OF THE ALGORITHM: The present research included a thorough investigation of the method of Figure 3, through both theoretical verification and numerical experiments that utilize measured data to investigate and verify the method. From a theoretical point of view, the derivation of Patience was found to be sound, except for one assumption that the term Z_{cor} (see Roth and Dahlke [6]) was assumed to be equal to the network parameter Z_{11} . This assumption is believed to be an unnecessary constraint.

Next, the Patience method was investigated on the basis of a set of numerical experiments made using measured data provided from industry for a MESFET. The measured data, the output from a commercially available noise parameter measurement system, consisted of measured noise figures with corresponding values of Γ_s at 12 frequencies, for 16 different source states per frequency. The algorithm (Figure 3) was then applied in two different ways to generate the data shown in Figure 4.

Figure 4a shows the results by applying the method using, for each source state, the noise figure data and Γ_s values provided versus frequency. Note that for perfect data all of these curves would be flat. The fact that they are not is due in part to errors in the measurement data. For example, the data corresponding to source state 2, 9, and 16 of Figure 4a appears questionable since C becomes greater than unity⁴. This illustrates how the method may be used to complement a noise parameter measurement procedure, by identifying problem source states. Some deviations may also be due to errors in the small signal model.

In Figure 4b are shown the results of applying a slightly modified algorithm at each frequency that uses the measured noise figures and Γ_s values provided as a function of source state. With accurate data, the experiment of Figure 4b can be used to investigate the frequency independence assumption for P, R, and C. Although most reported noise modeling methods have required or assumed P, R, and C to be frequency independent. Some numerical work has suggested that they may be frequency dependent in general [15-18]. This is a subject of continuing controversy^{5,6}, and more data needs to be examined before reaching any definitive conclusion.

Despite the variations in the P, R, and C values observed in Figure 4, the results for noise parameters determined from the algorithm of Figure 3 agree very well with measured noise parameter measurements as shown in Figure 5. This shows considerable promise for the method's application to develop mm-wave noise models with the improvements discussed next.

METHOD IMPROVEMENTS AND EXTENSIONS: Based on this research several improvements to the method have been initiated. These include the derivation of a more general equation for the noise figure that does not assume $Z_{cor} = Z_{11}$. The results of this derivation have now been implemented into the new method and verified with related software⁷. Other improvements initiated are the replacement of the synthesized "measurement" step from the algorithm of Figure 3 with a direct calculation of noise parameters, and the ability to solve for the noise source coefficients in different ways depending on the available measurement data. These ways include the use of multiple source state data to calculate P, R, and C (as implemented in Figure 4b), the use of F_{min} data versus frequency (similar to Robertson et. al. [28-29]), and the use of measured noise parameters at a single frequency or at multiple frequencies. This last set of data is the output of a commercial noise parameter measurement system. It follows that a physically based model, such as the present one, fitted to the entire set of available data may provide the most accurate frequency extrapolatable noise model.

Other issues to be considered in future research are the implementation of an improved small signal model extraction algorithm, and the application of the method to temperature dependent FET noise model development. The derivation of a small signal model, as done for the present work, from a single set of S-parameters at the operating bias, may not adequately separate the extrinsic and intrinsic model parameters. To address this problem, several methods have been proposed to determine a more accurate, physically meaningful, small signal FET model [33-34]. The University of South Florida is also investigating this problem as part of a separate project. In future

developments of this noise modeling method an improved small signal model procedure should be employed. Another area of strong interest to both the industry and the Air Force is the development of temperature dependent noise models for use in CAD and for use in predicting and explaining device noise behavior at temperatures from cryogenic (e.g. as low as 4K) to elevated temperatures (e.g. 475K). Consideration should, therefore, also be given to temperature dependent noise model development in future research.

V. SUMMARY AND CONCLUSIONS

This research project has been very successful in furthering the understanding of mm-wave noise modeling methodology. All of the research goals have been satisfied.

After a detailed literature review, it is concluded that the most viable methods for mm-wave FET noise modeling use a physically based noise equivalent circuit. Conventional noise parameter measurements are complicated at mm-wave frequencies, by the difficulty of fabrication and characterization of a low loss variable impedance tuner. This summer's research has advanced a new method that overcomes this difficulty by using a physically based noise equivalent circuit whose internal noise sources are characterized by noise figure measurements, made over a range of frequencies without the need for a tuner. The new method, therefore, may be used as a stand alone mm-wave noise modeling tool. Alternatively, the method may be used to complement a tuner based noise parameter measurement procedure in the following two ways. First, it can be used as a means to examine the integrity of the noise data corresponding to a particular source impedance tuner state. Second, it can be

used as a means to use the available measured noise data from the tuner procedure to derive a convenient frequency extrapolatable model.

This method builds upon the work of a previous AFOSR summer research project. An investigation of this previous work produced promising comparisons to measured data, but also revealed that modifications are necessary for reliable mm-wave model development. A new computer program that incorporates some of these improvements has been developed and verified. Work towards implementing the remaining improvements has been initiated. It is strongly encouraged that additional research be performed to more fully develop and verify the method advanced here, and to broaden its range of applicability to include temperature dependent FET noise characterization.

REFERENCES

- [1] H. Haus ed., "IRE Standards on Methods of Measuring Noise in Linear Twoports, 1959," *Proc. IRE*, vol. 48, pp60-68, Jan. 1960.
- [2] R. Lane, "The Determination of Device Noise Parameters," *Proc. IEEE*, vol. 57, pp1461-1462, Aug. 1969.
- [3] V. Adamian, A. Uhler Jr., "A Novel Procedure for Receiver Noise Characterization," *IEEE Trans. on Instrumentation and Meas.*, pp 181-182, June 1973.
- [4] M. Sannino, "On the Determination of Device Noise and Gain Parameters," *Proc. IEEE*, vol. 67, pp1364-1366, Sept. 1979.
- [5] W. Patience, "A Simplified Method of Determining Noise Parameters of High Frequency MESFET's," 1989 USAF-AFOSR Summer Faculty Research Program Final Report, Sept. 6, 1989.
- [6] H. Roth and W. Dahlke, "Theory of Noisy Fourpoles," *Proc. IRE*, vol. 44, pp811-818, June 1956.
- [7] H. Haus et. al., IRE Subcommittee 7.9 on Noise, "Representation of Noise in Linear Twoports," *Proc. IRE*, vol. 48, pp69-74, 1960.
- [8] A. van der Ziel, "Thermal Noise in Field-Effect Transistors," *Proc. IRE*, pp1808-1812, Aug. 1962.
- [9] A. van der Ziel, "Gate Noise in Field Effect Transistors at Moderately High Frequencies," *Proc. IRE*, pp461-467, March 1963.
- [10] A. van der Ziel and J. Ero, "Small-Signal, High-Frequency Theory of Field-Effect Transistors," *IEEE Trans. Electron Devices*, vol. ED-11, pp128-135, 1964.
- [11] F. Klaassen, "High-Frequency Noise of the Junction Field Effect Transistor," *IEEE Trans. Electron Devices*, vol. ED-14, pp368-373, July 1967.

- [12] W. Baechtold, "Noise Behavior of Schottky Barrier Gate Field-Effect Transistors at Microwave Frequencies," *IEEE Trans. Electron Devices*, vol. ED-18, pp97-104, Feb. 1971.
- [13] W. Baechtold, "Noise Behavior of GaAs Field-Effect Transistors with Short Gate Lengths," *IEEE Trans. Electron Devices*, vol. ED-19, No. 5, pp674-680, May 1972
- [14] R. Pucel, H. Haus, and H. Statz, "Signal and Noise Properties of GaAs Microwave FET," in *Advances in Electronics and Electron Physics*, vol. 38, L. Morton, Ed. New York: Academic Press, 1975.
- [15] B. Camez, A. Cappy, R. Fauquembergue, E. Constant, and G. Salmer, "Noise Modeling in Submicrometer-Gate FET's," *IEEE Trans. Electron Devices*, vol. ED-23, pp784-789, July 1981.
- [16] A. Cappy and W. Heinrich, "High-Frequency FET Noise Performance: A New Approach," *IEEE Trans. Electron Devices*, vol. ED-36, pp403-409, Feb. 1989.
- [17] W. Heinrich, "High-Frequency MESFET Noise Modeling Including Distributed Effects," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-37, pp836-842, May 1989.
- [18] W. Heinrich, "Corrections to 'High Frequency MESFET Noise Modeling Including Distributed Effects,'" *IEEE Trans. Microwave Theory Tech.*, vol. MTT-38, pp96-97, Jan. 1990.
- [19] T.M. Brookes, "The Noise Properties of High Electron Mobility Transistors," *IEEE Trans. Electron Devices*, vol. ED-33, pp52-57, Jan. 1986.
- [20] H. Fukui, "Optimal Noise Figure of Microwave GaAs MESFETs," *IEEE Trans. Electron Devices*, vol. ED-26, pp1032-1037, July 1979.
- [21] H. Fukui, "Design of Microwave GaAs MESFET's for Broad-Band Low-Noise Amplifiers," *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-27, pp643-650, July 1979.
- [22] H. Fukui, "Addendum to 'Design of Microwave GaAs MESFET's for Broad-Band Low-Noise Amplifiers,'" *IEEE Trans. Microwave Theory and Tech.*, vol. MTT-29, Oct. 1981.
- [23] A. Podell "A Functional GaAs FET Noise Model," *IEEE Trans. Electron Devices*, vol. ED-28, pp511-517, May 1981.
- [24] M.S. Gupta, O. Pitzalis, S. Rosenbaum, and P. Greiling, "Microwave Noise Characterization of GaAs MESFET's: Evaluation by On-Wafer Low-Frequency Output Noise Current Measurement," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, pp1208-1217, Dec. 1987.
- [25] M.S. Gupta, and P.T. Greiling, "Microwave Noise Characterization of GaAs MESFET's: Determination of Extrinsic Noise Parameters," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, pp745-751, April 1988.
- [26] M. Pospieszalski, "Modeling of Noise Parameters of MESFET's and MODFET's and Their Frequency and Temperature Dependence," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-37, pp1340-1350, Sept. 1989.
- [27] M. Pospieszalski, "FET Noise Model and On-Wafer Measurement of Noise Parameters," *1991 IEEE International MTT-S Symposium Dig.*, pp1117-1120, June 1991.
- [28] R. Robertson and T. Ha, "Noise Models for Gallium Arsenide Field-Effect Transistors at Room and Cryogenic Temperatures," *Int. J. Electronics*, vol. 61, No. 4, pp443-440, 1986.
- [29] R. Robertson and T. Ha, "Optimum Noise Source Impedance Determination for GaAs FETs at Room and Cryogenic Temperatures" *Int. J. Electronics*, vol. 63, No. 3, pp359-369, 1987.
- [30] R. Robertson and R. Sanders, "A Simplified Approach to Optimum Noise Source Impedance Determination for GaAs FET Amplifiers" *Int. J. Electronics*, vol. 65, No. 5, pp943-952, 1988.
- [31] A. Riddle "Extraction of FET Model Noise-Parameters from Measurement," *1991 IEEE International MTT-S Symposium Dig.*, pp1113-1116, June 1991.
- [32] L.P. Dunleavy, "A Ka-Band On-wafer S-parameter and Noise Figure Measurement System", *34th ARFTG Conference Digest*, December 1989.
- [33] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A New Method for Determining the FET Small-Signal Equivalent Circuit," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-36, No. 7, July 1988.
- [34] M. Berroth and R. Bosch, "Broad-Band Determination of the FET Small-Signal Equivalent Circuit," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-38, No. 7, July 1990.

FOOTNOTES

¹Equation (1) is also expressible in terms of source impedance or in terms of source reflection coefficient, in which case the optimum source impedance and reflection coefficient, respectively, are the parameters of interest.

² Personal communication with R. Webster, Rome Laboratory, April 8, 1991.

³See also "R&D Record Notebooks" of L. Dunleavy, and S. Lardizabal. A separate report is also being prepared to detail the results of the numerical experiments.

⁴ The noise sources, on a physical basis, cannot be more than 100% correlated.

⁵ Personal communication with R. Pucel, A. Riddle, W. Heinrich at 1991 IEEE MTT-S conference.

⁶On another controversial point, Riddle [31] claims that $1/f$ noise produced a significant influence on his microwave noise parameter calculations. In subsequent conversations with L. Dunleavy, both R. Pucel and M. Pospieszalski find this surprising and somewhat questionable, as $1/f$ noise is widely believed to be entirely negligible at microwave frequencies. We may wish to re-address this question in future research.

⁷Results and comparisons are discussed in the separate report mentioned above detailing the numerical experiments.

EVOLUTION OF FET NOISE CHARACTERIZATION SOLUTIONS

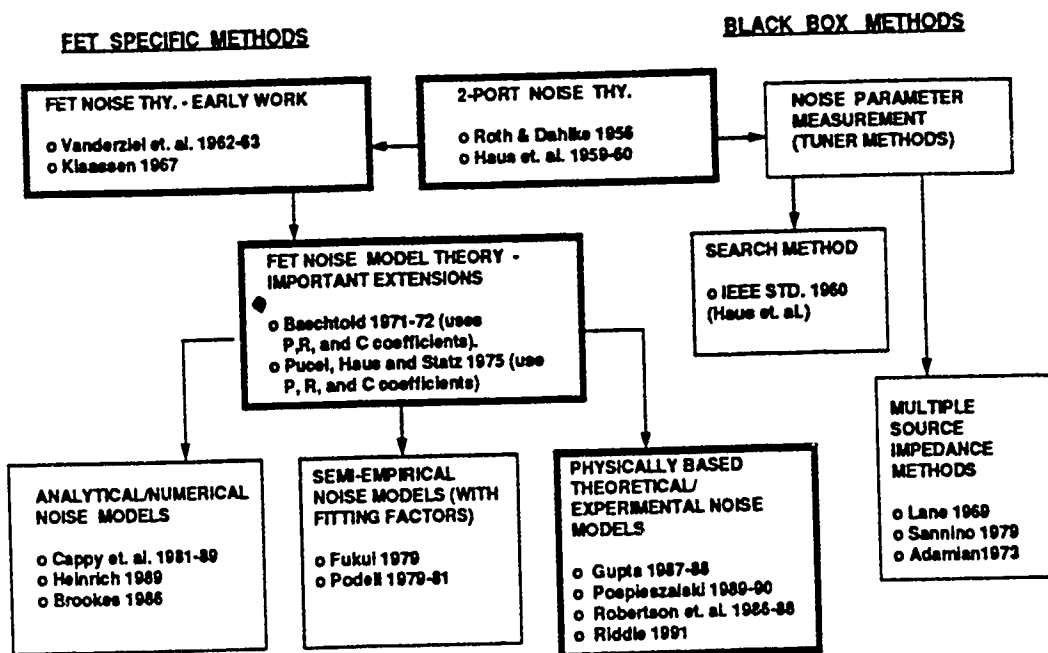
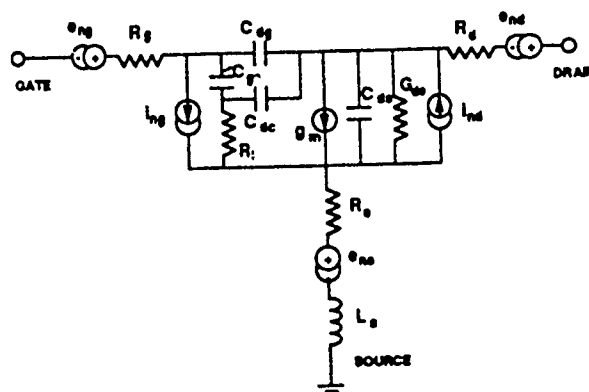


Figure 1. Flow chart showing the relationship between various reported noise characterization methods. All of the methods rely on the linear active two port noise theory developed by Roth and Dahlke.



| Noise Source Description | | |
|---|---|---|
| $\overline{e_{ng}^2} = 4kT\Delta R_g$ | $\overline{e_{nd}^2} = 4kT\Delta R_D$ | $\overline{e_{ng}^2} = 4kT\Delta R_g$ |
| $\overline{i_{ng}^2} = 4kT\Delta b_{ng}P$ | $\overline{i_{ng}^2} = \frac{4kT\Delta C_{gs}^2 \omega^2}{g_m} R$ | $\gamma_C = \frac{\overline{i_{ng}^2}}{\sqrt{\overline{i_{ng}^2} \overline{i_{ng}^2}}}$ |

Figure 2. Noise equivalent circuit used in the present research. Like similar circuits used by others, this circuit adds equivalent noise sources to a small signal FET model.

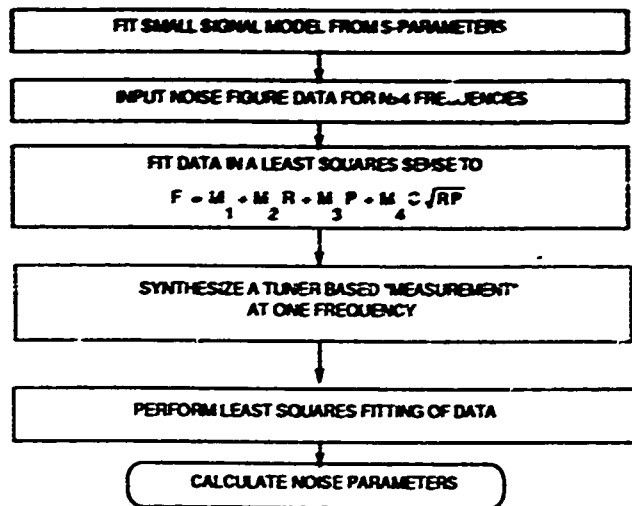


Figure 3. Algorithm for Noise Model Determination and Noise Parameter Calculation.

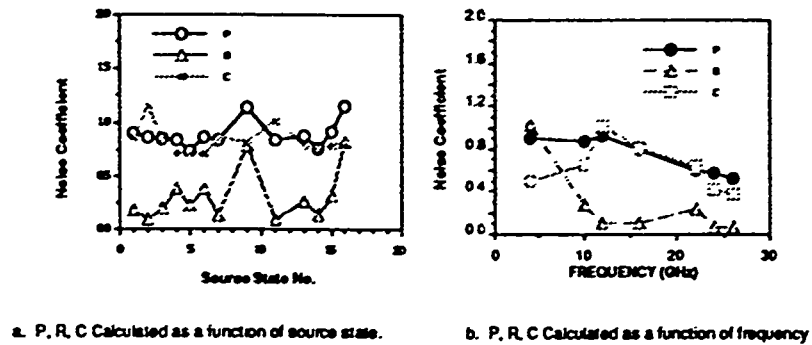


FIGURE 4. Results of numerical experiments performed with Patience's algorithm.

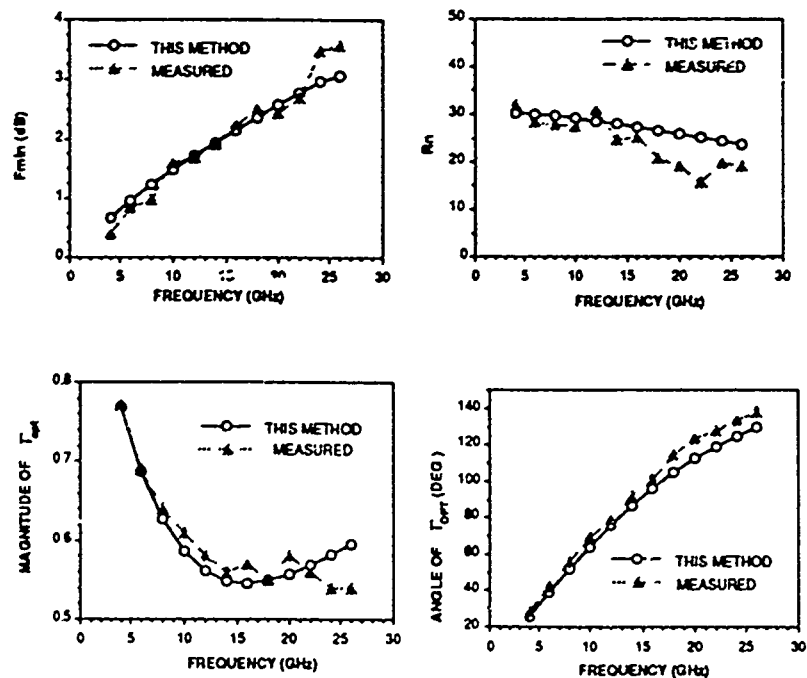


Figure 5. Comparison of noise parameters calculated for a MESFET using the present modeling method to measured noise parameters obtained with a tuner based procedure.

THOMAS LUSBY

REPORT NOT AVAILABLE
AT TIME OF
PUBLICATION

TESTINGS ON EFFECTS OF THERMAL TREATMENT ON GaAs (AND InP) SUBSTRATES AT MOCVD TEMPERATURE

Yonghuan Zhou

In MOCVD processing, substrates are developed at about 650°C. To minimize the thermal degradation effects at this temperature is crucial for effective quality control. A Heat Cell was designed earlier for investigating the optimum processing conditions. In order to test the protecting function of the heat cell, surface degradations of GaAs and InP substrates first without protection while being thermal treated was studied by photoluminescence method. Various experiment conditions including chemical treatments were examined to develop a procedure of effective and accurate measurement.

1. Introduction

Precise composition and uniformity of the substrates are essential for good performances of semiconductor devices. In MOCVD processing, substrates, such as GaAs and InP, are developed at about 650°C. At this temperature, thermal degradation effects are mainly responsible for surface decompositions and damages. Therefore, to characterize and minimize the thermal degradation effects is necessary for effective quality control in growing process.

A Heat Cell designed earlier will be used for *ex situ* investigation of the optimum processing conditions, namely temperature and ambient pressure, under which the thermal degradation will be minimized to below an acceptable level. It was designed to preserve ambient pressure of desired components during the thermal treatment. The aim of the experiment reported here is to examine the surface degradation of GaAs and InP substrates without protection while being heated. The photoluminescence method (PL) was used for measurements. The result will be compared to that in the case of with protection of the heat cell. (This part of experiment is to be completed later)

2. Experiments

GaAs and InP substrates were treated in a procedure consisting of chemical treatments and heating, and examined by photoluminescence method.

1. Samples : InP (n^- Sn doped); GaAs

2. chemical treatments:

Various procedures for surface cleaning, deoxidation and etching have been examined, including :

a) NRL etchant (1 HF, 1 HCl, 4 H_2O plus one drop of H_2O_2 per 12 ml solution) ^[1]

b) HIO_3 (10%), etching rate: 350 Å/min.

c) HF (1ml/l) ^[2]

d) HNO_3 ²

e) $H_3PO_4 : H_2O_2 : H_2O$ (3:1:50), etching rate: 600 Å/min. ^[3]

We tested solution a)-d) on InP samples. We found that NRL is effective to remove the surface oxide layer. The samples treated by NRL gave more PL intensities. But it has strong time dependence for a long time after the treatment. we chose solution e) for etching GaAs samples. The etching procedure is: immersing the sample in the solution for a certain time corresponding to certain etching depth, rinsing it with DI water, then dry it in air.

3. Heating:

Samples were heated under $N_2:H_2$ (90:5, 20 in Hg VAC.). The temperature was monitored by a thermocouple for below 400°C and a Vanzetti two-color pyrometer for above 400°C. After heating, samples were remained under $N_2:H_2$ until cooled to room temperature.

We also designed a compact heat plate that can be fit into a N_2 box with a quartz window on the top. So one can take the PL measurement with the sample set at a desired temperature.

4. Photoluminescence measurements:

Several techniques have been developed to determine the band structures and their dependence on composition of semiconduction materials, e.g. electroreflectance, spectroellipsometry etc. Among them, photoluminescence method is suitable for surface

characterization, because it is highly sensitive to crystal defects, thereby surface deformations. In addition, its whole optical nature provides high accuracy and nondestructivity which are desirable for both *in situ* controlling and *ex situ* testing.

The experimental setup for photoluminescence measurement is shown in Fig.1.

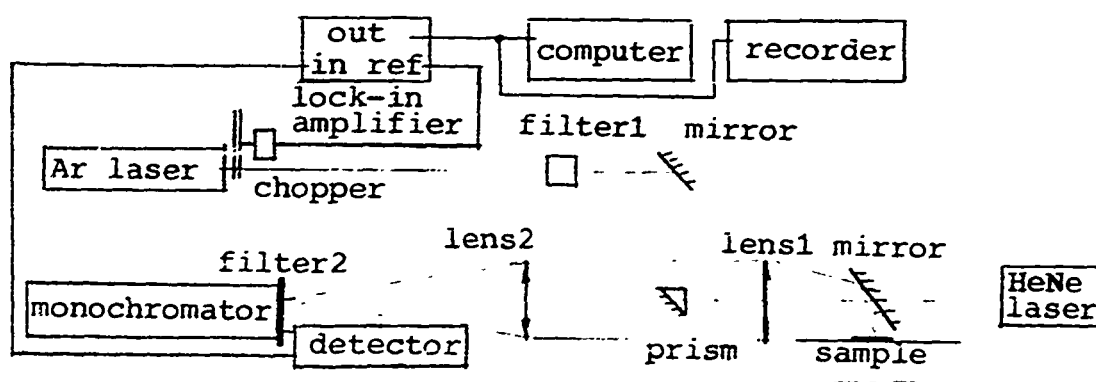


Fig.1 Experimental setup for PL measurements.

Filter 1 (blue) cuts off the lights other than 5145\AA . Filter 2 (red) allows the PL lights passing through which have peak intensities at about 8770\AA for GaAs and 9270\AA for *inP*. Samples were mounted on an X-Y translation stage. The HeNe laser is helpful for initial alignment.

The measurements were taken in two ways:

a) To focus the light at one point of the sample and take the PL spectrum over certain wavelength (energy) region.

b) To set the monochromator at the wavelength where the PL intensity peak is, and let the light scan over a line on the sample by moving the sample mount in one direction slowly while a recorder records the PL intensity. Later, to take the average value of PL intensity over the line. Since we have observed significant differences of PL intensities at different points of the sample, we have paid much attention to sample positioning. However, the peak shift during the measuring was negligible in this case, therefore method b) has the advantage of not only reducing the errors due to position changes, also giving some information about surface uniformity at each step of the procedure, and about time dependences of PL intensities after each treatment.

Laser power plays an important role for PL intensity. It is believed that at low level of power, PL intensity increases linearly with laser power. At high power level PL intensity

tends towards saturation.¹⁴ In order to avoid the nonlinear effect and the fluctuation of laser power, a calibration line of output power in terms of voltage (J4 reading) of the laser was measured. The laser power in the experiments remained constant. (35 mW at laser output).

We also compared the performances of three detectors:

- a) Ge detector, uncooled (room temperature);
- b) Ge detector, cooled (77°K) (North Coast);
- c) S1 detector(Photomultiplier tube)

We found satisfaction in the cooled Ge detector in regarding to both sensitivity and signal to noise ratio. The results showed in this report are all from using the cooled Ge detector.

3. Results

Three kinds of testing were accomplished by the experiments:

- a) Heating temperature:

The experiment results indicate that heating causes reduction in PL intensity for both GaAs and InP (Fig.3 and 5). It is especially severe for InP. Fig.4 and Fig.5 show that the depth of damage increases with heating temperature. At 600°C, the heating seems permanently reduced PL intensity for InP.

- b) Chemical treatment:

The fact that appropriate etching can bring the PL response, reduced by heating (Fig.2) or oxidation (Fig.3), back to original level by removing the damaged layer provides a convenient method for investigating the surface quality in terms of composition and uniformity. Our experiments show that $H_3PO_4:H_2O:H_2O$ solution is effective and reliable, suitable for our use.

- c) PL measurement:

The experiments also served as a testing on our setup for PL measurements. The result is that it meets our requirements in regarding to sensitivity, reliability and simplicity.

4. Conclusion

A reliable experiment setup for testing the effects of thermal treatment on semiconductor substrates has been accomplished. A meaningful experiment procedure was developed and proved. The experiment results gave the information about these effects, such as the scale of surface damaging, on GaAs and InP substrates in the case of without ambient protection. These results are valuable for they will be the witness of the protecting function of the Heat Cell which will be tested by the same experiment setup and same procedure.

References

- [1] R.R. Chang, R. Iyer, and D.L. Lile, J.Appl.Phys. **61**, 1995 (1987)
- [2] S.K Krawczyk and G. Hollinger, Appl.Phys.Lett. **45**, 870 (1984)
- [3] P.Viktorovitch, M. Gendry, S.K. Krawczyk, and F. Krafft, Appl.Phys.Lett. **58**, 2387 (1991)
- [4] C. Perry, private communication.

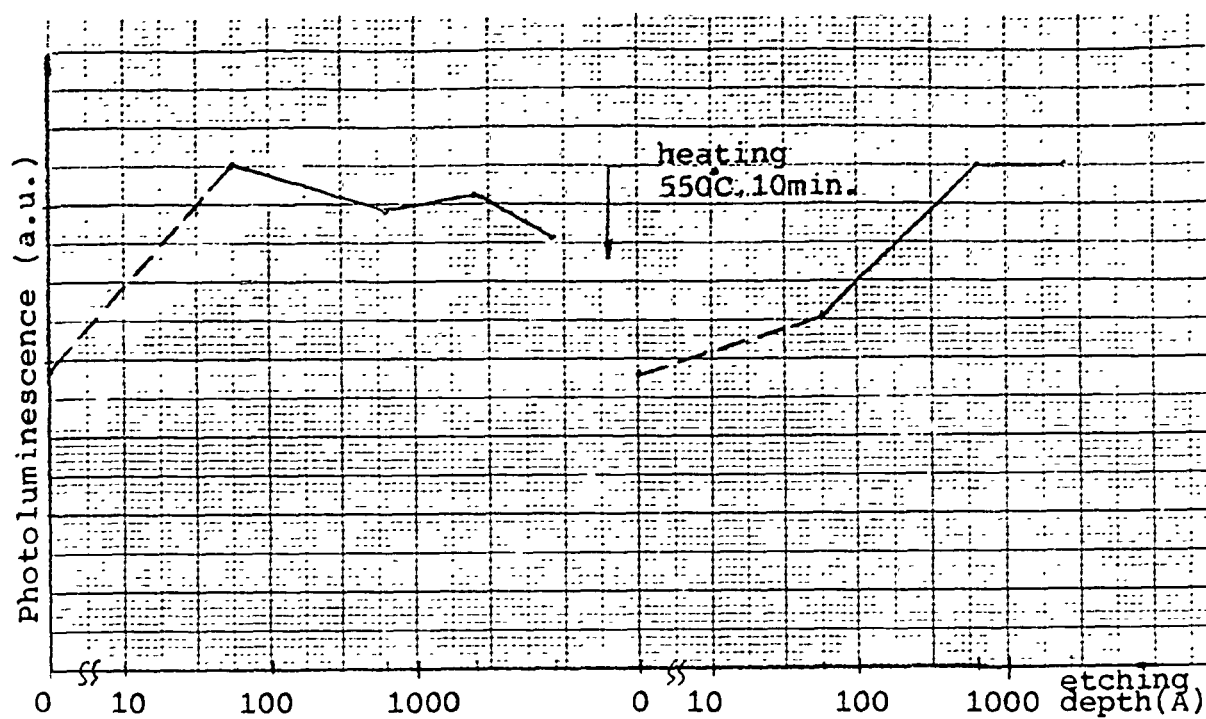


Fig.2 Photoluminescence intensity(average over a 1.3 mm line)of GaAs. Sample was etched, heated, then etched again. Etching solution: e (see text)

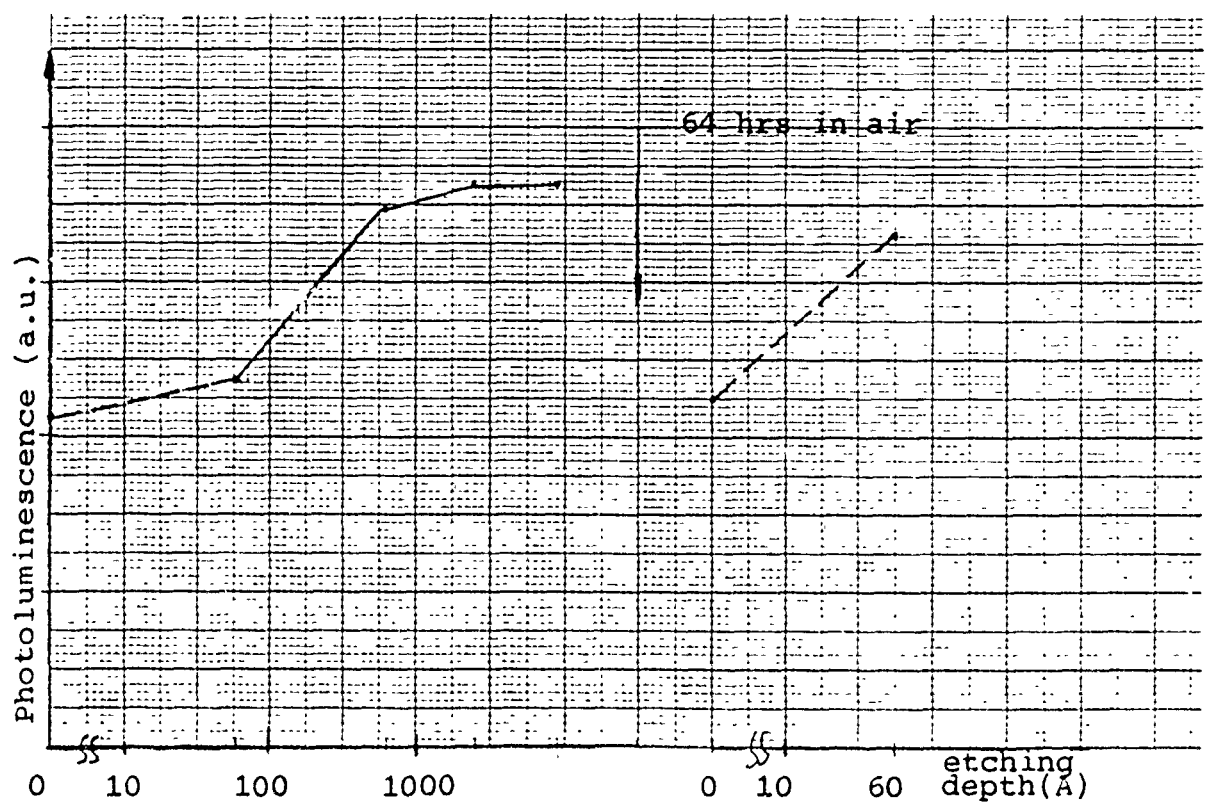


Fig.3 Photoluminescence intensity(average over a 1.3 mm line)of GaAs. Sample was heated at 600°C for 5 min. After etching, it was in the air for 64 hrs.

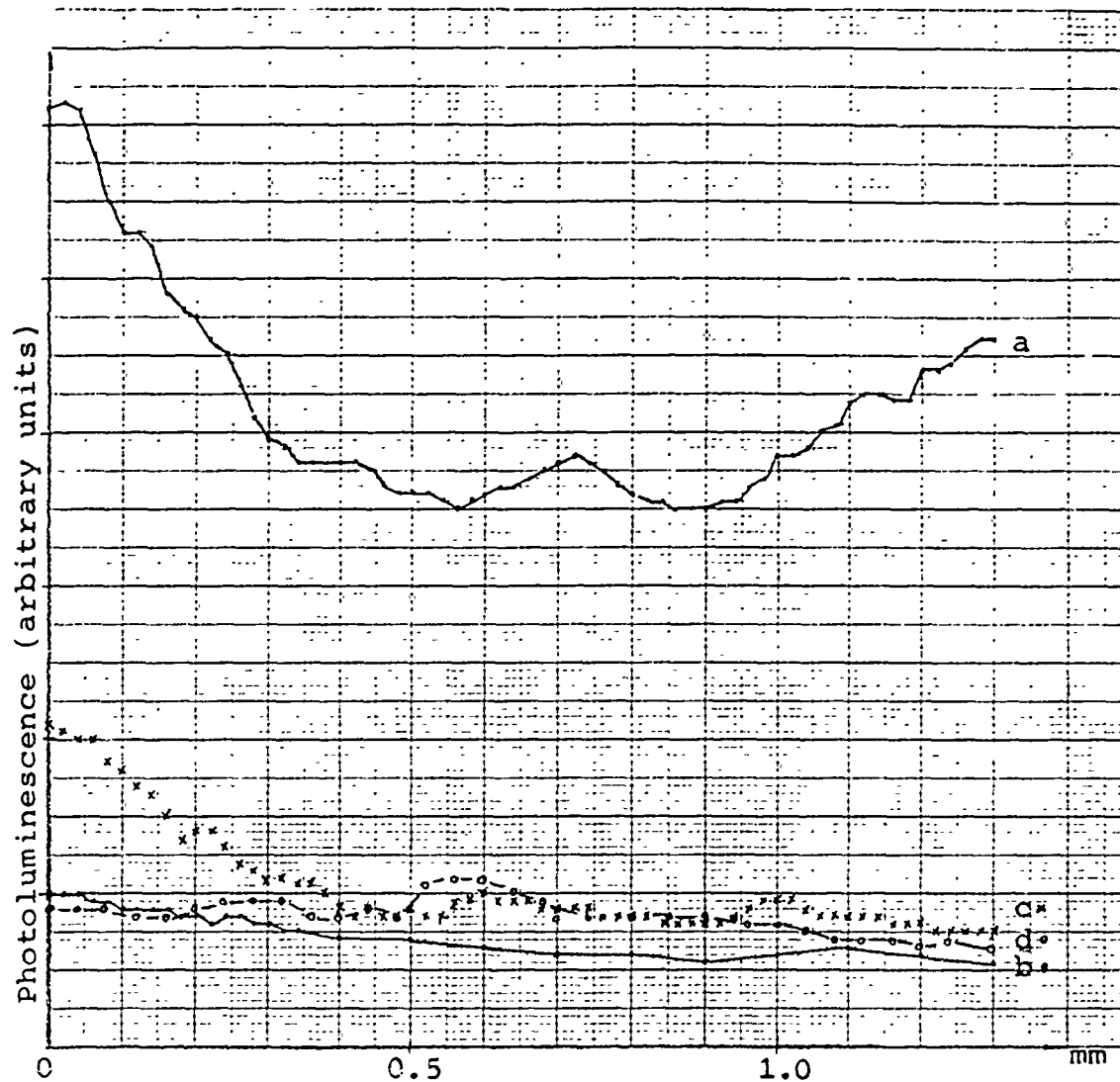


Fig.4 Photoluminescence intensity over a 1.3 mm line on InP sample. a) before heating, b)-d) after heating(5min., 600°C) and etching. b) 350 Å, c) 1050 Å, d) 4550 Å.

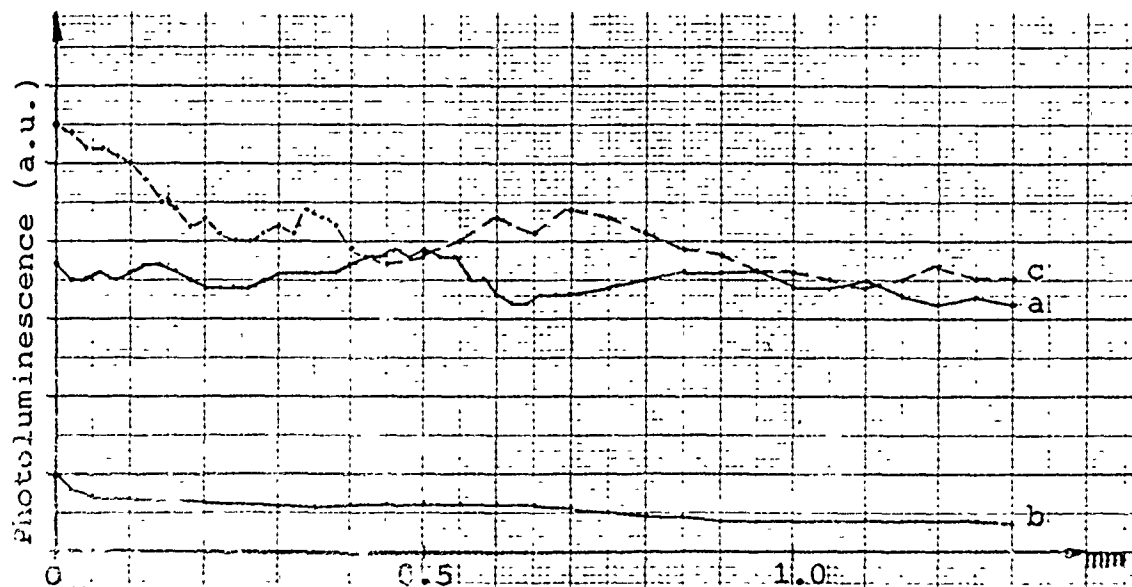


Fig.5 Photoluminescence intensity over a 1.3 mm line on InP sample. a) original b) after heating(5 min. 500°C) c) after etching 1200 Å.

NON-INTRUSIVE TESTING OF COMPOSITE
AIRCRAFT ENGINE COMPONENTS: II

James Abbey, Graduate Student
Engineering Science and Mechanics Program
Georgia Institute of Technology

ABSTRACT

This paper is the result of a ten week research effort to study non-intrusive testing of composite aircraft engine components. A review of polymer, metallic/intermetallic, and ceramic matrix composites is presented of which polymer matrix composites is concluded to be the most likely candidate. Also, various nondestructive testing methods as applied to composites are reviewed as well as new signal analysis methods such as acousto-ultrasonics and the wavelet transform. Then, the generation process of the promising NDT technique of laser ultrasonics is discussed in detail.

INTRODUCTION

The Air Force and the Sverdrup Corporation anticipate difficulty in non-intrusively evaluating the material condition of composite components. Anisotropy, inhomogeneity, and multiple failure mechanisms make evaluation

of these materials more challenging than for traditional metallic materials. Due to advances being made in composites research, components of these materials are expected to be utilized in the "next generation" of aircraft engines. New testing techniques need to be developed now so that the aerospace industry will be prepared when these advanced engines appear.

Both papers, part I and II, are the result of a ten week research effort funded by the AFOSR's Summer Faculty Research Program to address the testing problem. Part I, authored by Laurence Jacobs, states the problem and proposes the laser ultrasonics technique as a potential solution. Two aspects of laser ultrasonics, laser detection of ultrasound by interferometry and wave propagation in anisotropic materials, are then discussed in detail. This paper, part II, contains a review of composites, nondestructive testing techniques, and new signal analysis methods as relating to the problem. Also, a third aspect of laser ultrasonics, laser generation of ultrasound in solids, is discussed in detail.

COMPOSITES

Intensive research and development is currently being conducted through programs like NASA's High Temperature Engine Materials Technology Program (HITEMP) and Integrated High Performance Turbine Engine Technology Program (IHPTET) to bring the benefits of composite materials to the aerospace industry. This work is motivated by the high strength to

weight ratios that composites offer and their ability to be tailored for strength in specific directions. Both HITEMP and IHPTET are targeting the polymer matrix, metal matrix/intermetallic matrix, and the ceramic matrix composites.

Polymer matrix composites (PMCs) are the lightest of the three types of composites, but they have lower thermal oxidative stability than the others [1]. HITEMP has targeted the PMCs for an operating temperature of 425 degrees Celsius. This temperature has not yet been achieved, but notable progress has been made with the polyimides such as PMR-15 and Avimid-N [2,3,4]. These materials have the potential for turbine engine components such as air by-pass ducts, vanes, bearings, and nozzle flaps. In fact, the duct of the GE F-404 is made of PMR-15 with graphite fibers [4].

Most high temperature metallic matrix composites (MMCs) and intermetallic matrix composites (IMCs) have serious fiber/matrix compatibility problems [1,2,4,5]. One of these problems is mismatch of the coefficients of thermal expansion (CTE) of the fibers and the matrix. CTE mismatch creates residual stresses and may result in fiber/matrix breakage at high temperatures. MMCs and IMCs also typically have problems with excessive fiber/matrix reaction during fabrication and marginal high temperature oxidation resistance. Research has focused on the Aluminide matrices such as NiAl, Ni₃Al, FeAl, TiAl, Ti₃Al, and NbAl with SiC/Al₂O₃ fibers. New fibers need to be developed to overcome the current difficulties and reach

HITEMP's goal of a 1250 degree Celsius operating temperature [1].

The main problem facing ceramic matrix composites (CMCs) is the inability of the fibers to withstand the higher temperatures that the matrix can take. HITEMP wants the CMCs to withstand temperatures of up to 1650 degrees Celsius for gas turbine engines. Major focus is in developing small diameter fibers (<25 micrometers) to keep critical flaw size of these brittle materials as small as possible. CMCs also suffer from CTE mismatch and high temperature instability. As with the MMCs/IMCs, fiber development is the critical area of research [1].

In a report for the Sverdrup Corporation [4], Cruise concludes that the PMCs are the only "mature composites in terms of design properties and hardware experience." In light of this opinion, it is not surprising that our contact with Sverdrup, Tom Cromer, has informed us that the first composite they will be working with will probably be the polyimide Avimid-N from Du Pont.

NONDESTRUCTIVE TESTING TECHNIQUES

There are several different techniques to nondestructively evaluate a materials' condition. Many of these techniques have been considered for their potential application to composite materials.

Strain Gages

Strain gages are electrical devices used to measure strains from changes in resistance. The traditional, well-established method of measuring static and dynamic strains of engine components has utilized these devices. When applied to composite materials, though, the measured results can be misinterpreted due to fiber orientation and lamination. Only surface strain measurements with strong dependence on gage orientation can be obtained [6]. These problems coupled with the existing difficulties in high temperature mounting and telemetry or sliprings for moving parts make the technique less appealing for test engineers [7].

Laser Extensometry

A technique called laser extensometry [8] is quite simply an optical strain gage. This device employs lenses to focus two pairs of beams onto different points normal to a surface. The interference of the beams produce fringe patterns at each of the points and measure the relative displacement between the points. The technique has much potential because it is a non-contact technique. But for composites, it is also limited to a surface measurement like the strain gage.

Interference Techniques

Shearography, electronic speckle pattern interferometry (ESPI), holographic interferometry, and moire interferometry are all methods that use the changes in optical interference

patterns to access surface displacements in a full field [9,10,11,12]. These techniques are highly sensitive to changes and non-contacting, but they tend to be slow and yield only qualitative data. All these techniques have shown limited success in detecting flaws in composite materials.

Thermography

Thermography consists of imaging in the infrared portion of the electromagnetic spectrum. A thermally loaded object emits IR radiation with a spectral content and intensity related to its temperature and the nature of its surface. The technique has shown some promise for composites by detecting water inclusions and delaminations [13,14]. But, two drawbacks of thermography are its inability to discriminate between surface and subsurface features and its lack of sensitivity [13]. This technique will probably prove to be very useful in the future for detecting flaws in thin composites.

Radiography

A radiograph is like a shadow picture taken with x-rays. The darker regions of images correspond to the more penetrable areas of an object and the lighter ones to the more opaque areas. Radiography has been shown to be a very useful method for nondestructive testing of traditional materials by detecting voids, porosity, and inclusions [15]. Radiography has also been used to detect fiber fractures and fiber

orientations in thin composites [13]. But, the technique cannot discern the depth of these defects.

Computed Tomography

Computed tomography is very similar to radiography except that it images crosssections of objects. This allows CT to pinpoint defects such as porosity, voids, delaminations, and cracks very effectively [16]. The main problems with CT are its expense, slowness, and its need for calibration of x-ray attenuation coefficients to reveal physical properties such as density. Computed tomography will probably be used to determine the integrity of advanced components at the manufacturing level.

Traditional Ultrasonics

Traditional ultrasonics includes a range of NDT techniques in which stress waves are sent into an object via a transducer and couplant and then detected in the same manner after the stress wave passes through the object. The returning signals are characterized by attenuation, dispersion, and arrival times which all correspond to the object's properties and geometry. Ultrasonics has been successful in imaging, detecting defects, and measuring temperature variations, stress, and elastic constants in traditional metallic materials. In composites, measurements are more difficult because of anisotropic wave propagation,

but good measurements of the elastic constants have been obtained.

Laser Ultrasonics

Laser ultrasonics differs from traditional ultrasonics in the generation and detection process. Laser ultrasonics uses lasers to generate and detect stress waves, making it a non-contact, remote technique. Theoretically, the measurements made from traditional ultrasonics can all be made with laser ultrasonics. The technique has shown to yield good measurements of elastic constants and other properties in metallic materials, but more research needs to be conducted with respect to composite materials.

Eddy Current

The eddy current method uses close proximity transducers with high frequency currents to induce eddy currents in the surface of the tested object. These induced currents, in turn, affect the currents in the transducer which can be analyzed to reveal information concerning the surface of the object. In this way, eddy current testing is used to detect cracks, fiber breakage, fiber orientation, and delaminations in composites [17]. The method is highly sensitive, but is limited to electrically conductive materials/fibers at subsurface depths of 5 mm or less [18].

NEW SIGNAL ANALYSIS METHODS

Acousto-ultrasonics (AU) is a signal analysis method strongly linked to traditional ultrasonics. It consists of pulsing a transducer and receiving with another transducer some gage distance away. What makes the method special is that the receiver will count the number of peaks received above a preset voltage level. This number has been coined the stress wave factor by Alex Vary. It has been shown to be an empirical value that has a strong dependence upon geometry and attenuation. AU has been used to detect defects and measure strength qualitatively in composite materials [19]. The AU method also has the potential to be applied to laser ultrasonics.

Another signal analysis method that has been developed rather recently is the wavelet transform [20,21]. The wavelet transform is a mathematical transform of a one dimensional signal that can be used to detect the frequency content of certain areas in time. This method reveals time information where it is completely lost using the standard Fourier transform. This method is just starting to be applied and seems like it may be a powerful tool in the analysis of ultrasonic signals [20].

LASER GENERATION OF ULTRASOUND

Laser ultrasonics is a non-contact, remote technique that utilizes lasers to generate and detect ultrasound in solids. The ultrasound is generated by rapidly heating the solid's

surface. Generation occurs through two distinct processes of thermoelastic expansion and/or ablation. Detection of the ultrasound occurs by measuring the surface motion via laser interferometry. Part I of this report reviewed the detection process, while this part will cover the generation process.

When laser radiation is suddenly incident upon a solid's surface, some of the energy will be absorbed by the surface. The increase in energy within the material corresponds to an increase in temperature. When the incident radiation is of sufficiently low power densities as to avoid melting any of the material, the generation process is governed by thermoelastic expansion. The absorbed energy causes the material to expand rapidly and in turn creates a stress wave that propagates into the solid.

Force models in conjunction with wave propagation theory can be used for the prediction of ultrasonic displacements. A simple model can be constructed to represent the generation forces in the thermoelastic regime by assuming that the radiation is incident at a point and that thermal diffusion effects are ignored for short duration laser pulses. The model consists of two equal and perpendicular dipole forces acting at a point just below the surface and in the plane of the surface. This model is shown in Figure 1. These forces exhibit a step function time dependence. The model applied to a plate has shown that displacements are proportional to the energy of the laser pulse. The epicentral displacement for a plate has been calculated and is shown in Figure 2. The

arrival of the longitudinal wave corresponds to small, quick depression of the surface towards the source followed by a gradual depression of the surface. At the arrival of the shear wave, there is a larger step-like rise of the surface away from the source. Rose has shown for a plate that the ratio of the heights of the longitudinal step to the shear step is proportional to Poisson's ratio [22]. Experiments have shown this model to be valid for all of these theoretical features except that the experiments also show a small precursor at the arrival of the longitudinal wave [23]. This precursor has been attributed to thermal diffusion effects, ignored in this model, and optical penetration into the depth of the surface [23].

The model predicts that energy for the longitudinal waves will not be directed evenly into a half-space, but will have lobes directed at about 30 degrees to the surface. Experiments have also shown good agreement with these results. This type of graph is known as an energy directivity pattern.

When higher power densities of laser energy are used, melted material ejects from the object's surface. This ejected material causes a recoil force, according the conservation of momentum, that causes a stress wave to propagate into the surface. This process has also been modeled with the same assumptions as the thermoelastic model except that this one consists of a normal force at a point with a delta function time dependence [23]. See Figure 3. When applied to a half-space, the epicentral displacement

corresponds to a positive pulse at the arrival of the longitudinal wave with a small change in slope at the arrival of the shear wave. See Figure 4. Experiment has also shown this model to be valid for the half-space. The energy directivity pattern is fairly uniform in all directions for the longitudinal waves in theory and experiment.

CONCLUSION

A review of various nondestructive testing techniques led to a decision to chose the laser ultrasonics method for study. Laser ultrasonics is a sound choice because of the success of traditional ultrasonics and the promise that laser ultrasonics has in bringing this success to a non-contact world. Laser ultrasonics may be the method needed to access the material condition of advanced composites like Avimid-N.

APPENDIX

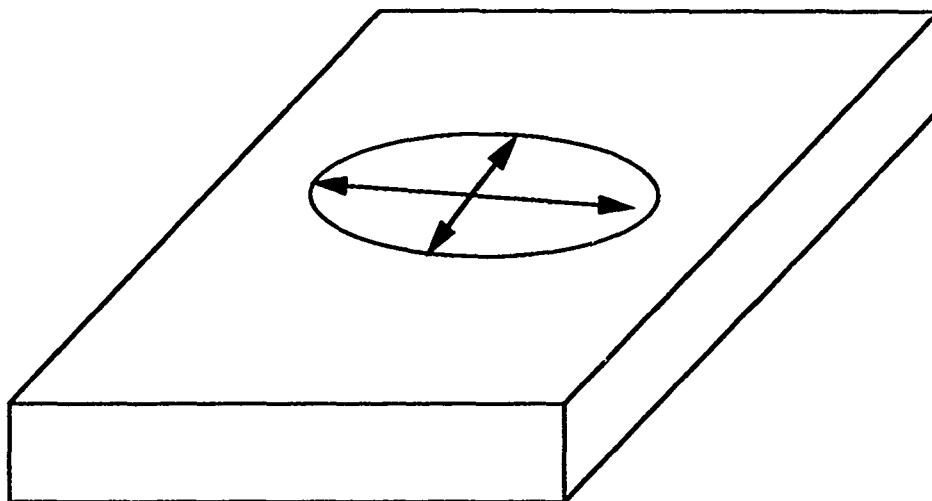


Figure 1.

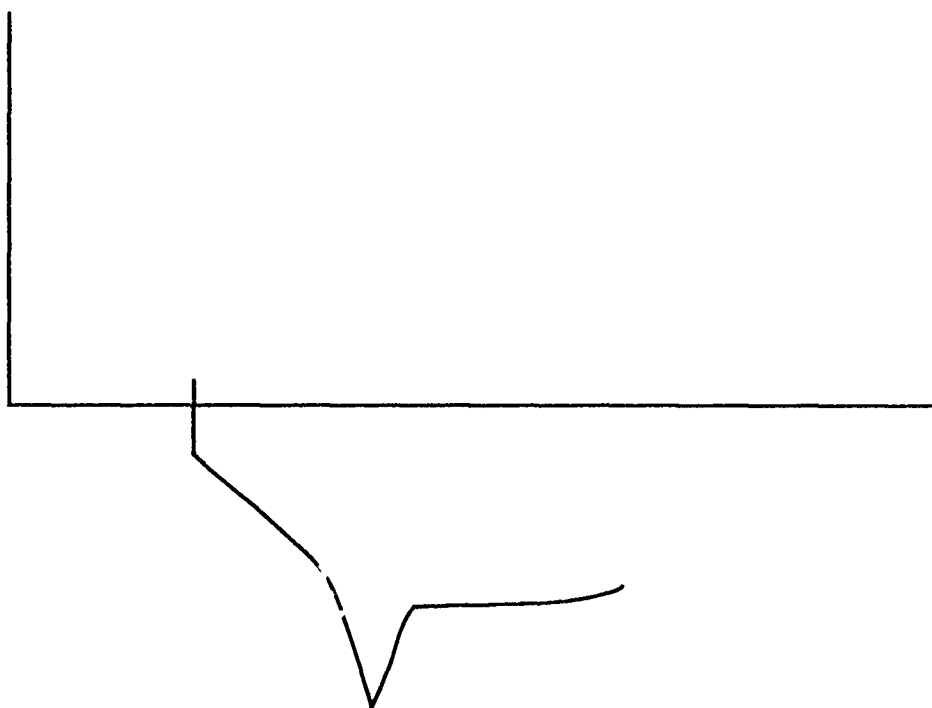


Figure 2

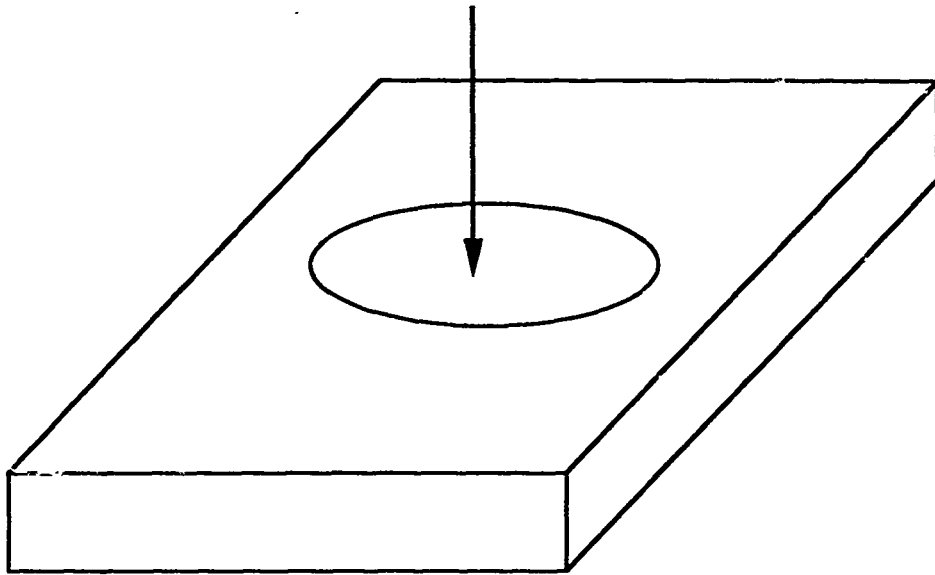


Figure 3.

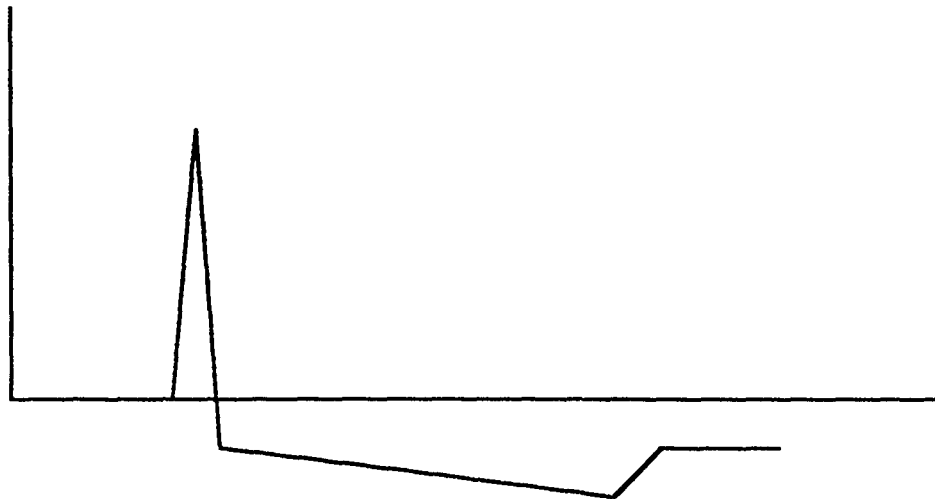


Figure 4.

BIBLIOGRAPHY

1. Stephens, J.R., "Composites Boost 21st Century Aircraft Engines," Advanced Materials and Processes, April 1990 pp. 35-38.
2. Gray, H.R., "Advanced High Temperature Engine Materials Technology Program," Aeropropulsion 1991, NASA-CP-10063 (N91-20086), pp. 24-28.
3. Scola, D.A., "The Status of High Temperature Polymers for Composites - Likely Composites," 34th International SAMPE Symposium, May 8-11, 1989, pp. 246-258.
4. Cruise, T.A., "Composite Materials for Aeropropulsion and Space Propulsion Structures," Final Report for the Sverdrup Corporation, June 30, 1991.
5. Sims, C.T., "Non-Metallic Materials for Gas Turbine Engines," Advanced Materials and Processes, June 1991, pp. 32-39.
6. Tuttle, M.E. and Brinson, H.F., "Resistance-foil Strain-gage Technology as Applied to Composite Materials," Experimental Mechanics, March 1984, pp. 54-65.
7. Kadmabi, J.R., Quinn, R.D., and Adams, M.L., "Turbomachinery Blade Vibration And Dynamic Stress Measurement Utilizing Nonintrusive Techniques," ASME Journal of Turbomachinery, Vol. III, October 1989, pp. 468-474.
8. Creighton, A.M. and Hercher, M., "Zeeman-split-laser System Measures Material Strains," Laser Focus World, June 1991, pp. 87-89.
9. Vikhagen, E. and O.J. Lokberg, "Detection of Defects in Composite Materials by Television Holography and Image Processing," Materials Evaluation, Vol. 48, February 1990, pp. 244-248.
10. Lin, C.H., Genge G.G., Pearce, L.J., Suhling, J.C., and Turner, J.L., "Evaluation of Damage in Composite Materials Using Digital Comparative Holography," SEM Spring Conference on Experimental Mechanics, 1989, pp. 18-26.

11. Czarnek, R., Post, D., and Guo Y., "Nonuniformities In Composite Panels by Moire Interferometry," Proceedings of the 1986 SEM Spring Conference on Experimental Mechanics, New Orleans, June 8-13, pp. 630-635.
12. Anastasi, R.F., Serabian, S.M., Shuford, R.J., and Das-Gupta, D.K., "Nondestructive Detection of Simulated Delaminations in Composite Laminates by Laser-Speckle Shearography," Experimental Mechanics, June 1987, pp. 28-31.
13. Leonard, L., "Testing for Damage," Advanced Composites, Jan./Feb. 1991, pp. 40-45.
14. Connolly M. and Copley, D., "Thermographic Inspection of Composite Materials," Materials Evaluation, December 1990, pp. 1461-1463.
15. Bray and Stanley, "Nondestructive Evaluation: A Tool for Design, Manufacturing, and Service," McGraw-Hill, New York, 1989.
16. Hanson, D.B., "Computed Tomography Imaging of Advanced Gas Turbine Engine Components", October 1990, WRDC-TR-90-2110.
17. Valleau, A.R., "Eddy Current Nondestructive Testing of Graphite Composite Materials," Materials Evaluation, Vol. 48, February 1990, pp. 230-239.
18. Sharpe R.S., Research Techniques in Nondestructive Testing, Vol. 7, 1985, Academic Press, London.
19. Duke, Jr. J.C., Henneke II, E.G., Kiernan, M.T., and Grosskopf, P.P., "A Study of the Stress Wave Factor for Evaluation of Composite Materials," NASA Contractor Report 4195, January 1989.
20. David, P.M. and Chapron, B., "Underwater Acoustic Signal Analysis with Wavelet Process," JASA 87(5), May 1990, pp. 2118-2121.
21. Kumar, A. and Fuhrmann, D.R., "The Frazier-Jawerth Transform," 1990 IEEE International Conference on Acoustic, Speech, and Signal Processing, Vol. 5, pp. 2483-2486.
22. Rose, L.R.F., "Point-source Representation for Laser-generated Ultrasound," JASA 75(3), March 1984, pp. 723-732.
23. Hutchins, D.A., "Ultrasonic Generation by Pulsed Lasers," Physical Acoustics, Academic Press, 1988, pp. 21-123.

PARALLEL SIGNAL PROCESSING FOR TURBINE ENGINE TESTING

| | |
|----------------|--------------------------------------|
| Prepared by: | Theodore A. Bapty & Ben A. Abbott |
| Academic Rank: | Graduate Students |
| Department: | Department of Electrical Engineering |
| University: | Vanderbilt University |

ABSTRACT

Arnold Engineering Development is the preeminent facility for ground testing of turbine engines at simulated altitude. The AEDC test facilities can duplicate the conditions experienced by the engine during an actual flight in a controlled environment. Consequently, the testing process is much safer and less costly than the alternate method of flight testing.

In order to assess the performance of the engine, they are heavily instrumented with various sensors. The outputs of sensors are sampled at high rates to capture information about high frequency processes occurring inside the engine. These outputs must be processed further to reveal the information. These signal processing computations are typically computationally intensive. The Fast Fourier Transform, for instance, requires approximately 34 floating point computations per sample under normal conditions. Immediate feedback is beneficial to the engineers testing the mechanical system. Online results enable the engineers to make interactive decisions during the test, reducing the cost and/or improving the quality of the data. The combination of quick turnaround and high computation per channel impose large computational requirements on any system designed to solve these problems.

The high computational requirements and natural parallelism make parallel processing an effective approach. A transputer-based, heterogeneous parallel processing system has been designed to meet the system goals. A software methodology has been developed to ease the programming tasks and to increase the flexibility of application software.

This report details the results of the Graduate Summer Research Program in the development of a system for the on-line analysis of these signals. The system has been under development at AEDC for several years. The results of the summer research have been to define a hardware/software architecture to enable the system to be scaled to a large number of channels. Another key issue addressed during the summer has been the demonstration of system capabilities to potential users for feedback and requirements review. The predominant response has been extremely positive.

I. INTRODUCTION:

Arnold Engineering Development is the preeminent facility for ground testing of turbine engines at simulated altitude. The AEDC test facilities can duplicate the conditions experienced by the engine during an actual flight in a controlled environment. Parameters such as air flow, pressure (simulated altitude), temperature and engine throttle settings can be varied as desired to create the necessary simulated conditions. By varying these and other parameters, the full flight envelope can be presented to the engine. Since all of the tests are carried out on the ground, the testing process is much safer and less costly than the alternate method of flight testing.

When testing turbine engines, an engine is installed in a ground test cell to simulate a wide variety of conditions. Stress testing is a specific type of testing used to analyze the aeromechanical performance of an engine. To analyze the dynamic vibrations, stress sensors are placed on the turbine fan blades and stators. A typical stress test instruments the engine with several hundred of these stress sensors along with a variety of temperature, pressure, flow and revolution per minute sensors(Figure 1). Stress sensors can generate signals with bandwidths in the tens of kilohertz.

The turbine engine is heavily instrumented with various sensors. The outputs of sensors

are sampled at high rates to capture information about high frequency processes occurring inside the engine. The outputs must be processed further to reveal the information contained therein. These signal processing operations are typically computationally intensive. The Fast Fourier Transform, for instance, requires approximately 34 floating point computations per sample under normal conditions. Thus to perform the basic translation to frequency domain results in an extremely heavy load. For example, on a test of 100 channels, for 100 seconds, sampled at 100 KHz demands a total of

$$100 * 100 * 100,000 * 34 = 34 \text{ billion}$$

floating point computations on a dataset of 1 billion numbers. These computations are only part of the algorithm used in the analysis of the data.

Immediate feedback is extremely beneficial to the engineers testing the mechanical system. Online results enable the engineers to make interactive decisions during the test. By making these educated decisions, the cost of the test can be reduced. Alternately, the information can be used to complete the dataset in the region of "interesting" conditions, thus improving the quality and completeness of the data set. In addition, the data may reveal pathological conditions which can result in an system failure. Seeing these results online may help avoid the loss of an expensive prototype and reduce the risk to test facilities and personnel.

The combination of quick turnaround and high number of operations per channel impose large computational requirements on any system designed to solve these problems. In the example above, spreading the computations out over the 100 second test period, an aggregate computation rate of 340 MFLOPS must be *sustained*. The estimates for future tests, when combined with the other processing requirements, top *two GFLOPS sustained*.

Table 1 lists performance numbers of some commercially available computers. These performance figures are from the NAS benchmarks [4] as well as experiments. The table shows that the high end of the world's fastest supercomputers cannot handle even the limited case of the example. Furthermore, the FFT is well suited to the vector processing architecture of the commercial supercomputers. Other algorithms used in the test data analysis are significantly less well suited to these architectures.

More problems become apparent when the requirements are analyzed from the perspective of bandwidth. Again, using the previous example, the *input* bandwidth is 10

TABLE I
MFLOPS RATINGS FOR 1024 POINT FFT

| Computer System | Sustained MFLOPS |
|------------------------------|------------------|
| Cray Y/MP (optimized) | 160 |
| IBM PC/386 | 0.7 |
| T800-20/Logical Systems C | 0.6 |
| T800-20 + VSP (Local Result) | 38 |

Megasamples per second, or 20 MBytes per second. As the data is translated into floating point, and transformed into a frequency representation, the data volume increases. The internal system bandwidth for this simple example can easily expand to several hundred megabytes per second simply to transfer data around within the system. This does not include the bandwidth required internal to the computations. Note that this bandwidth is required in addition to, and concurrently with the computational load. This is an area which typically imposes heavily upon the vector supercomputer architecture.

The data acquired during testing must be recoverable for further processing and analysis. Consequently the information must be stored in a format which allows full access to its information. This requirement, in its minimal form, mandates fully recoverable storage of all input signals at full bandwidth. As we have seen, this requires rates in excess of tens of megabytes per second. When the testing periods last for 5-12 hours, at these rates the total storage mounts into the Terrabytes. To minimize the storage media, some form of compression is needed, and the data should be stored in its most compact form.

In order to convey the results of the online data processing, a high-speed display device is required. Typically, the graphical format is one of the more efficient methods for presentation of results. Standard engineering graphics show the results in a manner which is familiar to the test engineers. Adequate resolution is needed to display the multitude of sensors. Updates must occur smoothly, to avoid operator fatigue, and to minimize the chance of missing a physical engine event.

The high computational requirements and natural parallelism make parallel processing a cost-effective hardware approach. Due to the extremely high computation requirements and

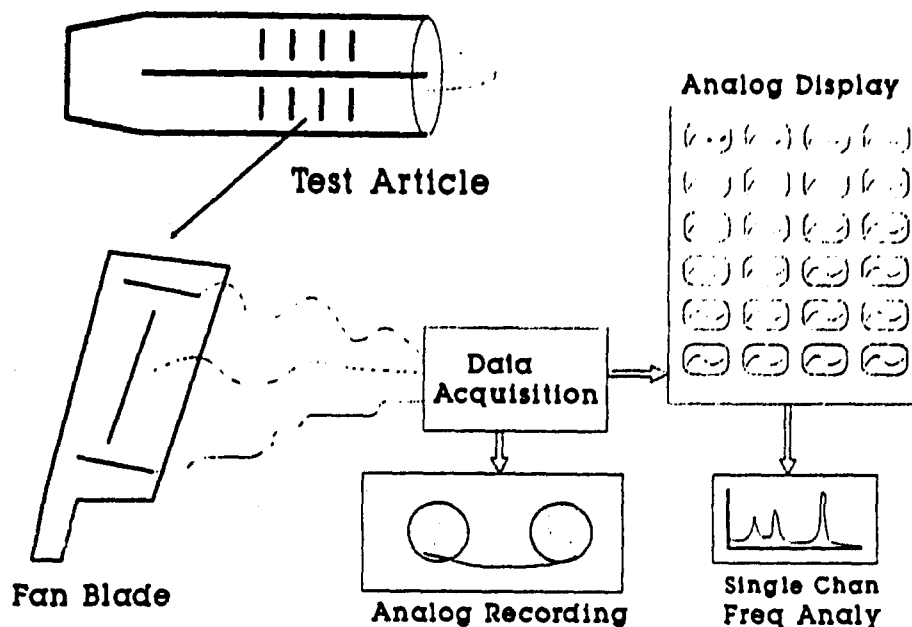


Figure 1: Turbine Engine Instrumentation

data bandwidths, a distributed memory architecture is most appropriate. In addition, the wide variety of requirements (MFLOPS, Storage, Display, I/O) imply that a heterogeneous architecture is needed.

To achieve the multi-GFLOPS goal, a large number of processors must be applied. The distributed memory Transputer architecture matches most of these requirements. The bare transputer however, lacks the computational capabilities required. At 2 MFLOPS (Peak) execution speed, over 1000 processors would be needed to achieve the 2 GFLOPS requirement. The intermediate storage of results exceed the reasonable capacity for semiconductor memory. Also, the bare transputer has no graphics output or mass storage capability. The approach with this effort has been to modify the transputer architecture with special-purpose devices to adapt the processing element to a specific task. What follows is a description of the project status as of the beginning of the summer.

Project History

The transputer-based, prototype heterogeneous parallel processing system has been designed to meet the system goals. While the full scale (24 channel) system has not yet been

implemented, all processor modules are functional:

1. **The Numeric Processing Element** uses a **T800** and a **Zoran floating point Vector Signal Processor** [10] to supply the bulk of the computational horsepower.
2. **The Graphics Processing Element** incorporates a **Texas Instruments TMS34010** [9] with a 1024x768, 16 color, dual-page frame buffer. This processing element performs quick, smooth display of dynamic signals. This processing element runs **GNU-Plot**, a standard engineering graphics package with a high-level command interface, with modifications to plot on-line data streams. Display rates range up to 20 updates per second.
3. **The Storage Processing Element** uses a DMA-equipped **SCSI** interface to store large quantities of information. An extremely simple statically allocated, flat file system has been implemented to provide efficient, albeit primitive, data storage.
4. **The I/O Processing Element** provides a variety of **RS233**, **Centronics**, **IEEE-488** and configurable digital parallel I/O lines. This processing element interfaces with the **Front End Processor**, and with printers, plotters, and commercially available instrumentation.
5. **The Front End Processor** performs the **A/D** conversion, and uses a three **Motorola 56001 Digital Signal Processors** [6] to complete the high-rate peak detection, **RMS** validation, and time & frequency domain alarm-level monitoring. The **FEP** also digitally antialias filters the data down to the system's operational bandwidth.

A **Transputer** is used in conjunction with each of the processing nodes. It provides the general message passing and control fabric of the system. A demonstration system has been developed to process two channels of high bandwidth data, and four channels of low bandwidth data. The block diagram of this system is displayed in figure 3. This system has the capability to:

1. **Store** all of the low bandwidth signals, and, at a lower rate, store processed data to a **DOS** file system,

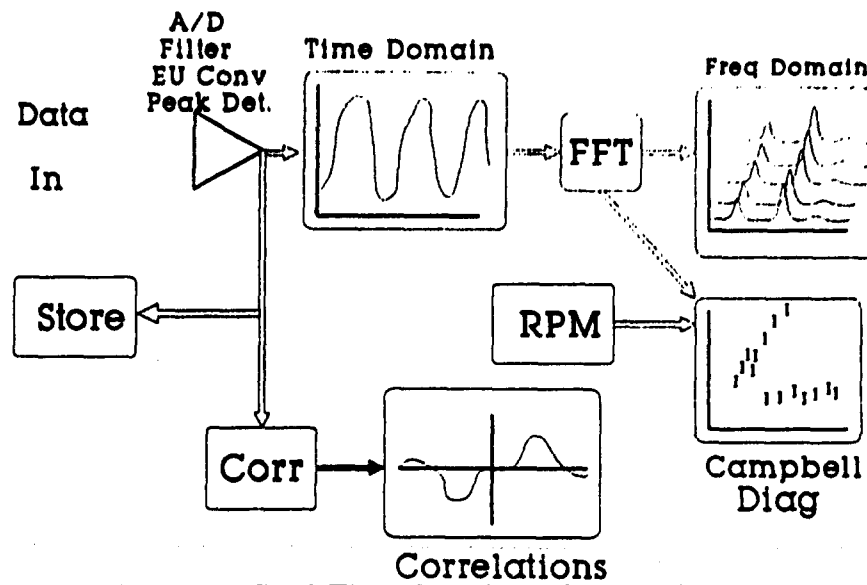


Figure 2: Real-Time Data Visualization System

2. **Monitor** continuously all input signals for excessive stress,
3. **Process** continuously all input signals for generation of time, frequency, crosscorrelation, and various types of Campbell diagrams,
4. **Display** results continuously, with up to 16 windows concurrently updating. With smaller numbers of windows, updates approach 20 per second. and,
5. **Interact** with the user to allow specification of screen windows and the graphics appearing on the screen.

Summer Research Focus

The state-of-the-art in software development for large scale parallel processing is still quite primitive, especially in time-critical applications such as mechanical system testing. Programming these systems with standard software engineering techniques, as developed for mono-processor systems, is labor-intensive and error-prone. This process is complicated even further by the requirements for these systems to be dynamically changing to satisfy the user's demands.

In an effort to confront these software development difficulties, Vanderbilt researchers have developed the Multigraph Kernel (MGK) [1]. The MGK provides a generic execution environment for parallel programs built by various tools comprising the Multigraph Programming Environment (MPE). The environment can be viewed on several levels. At the bottom is the Multigraph Kernel, or MGK. The kernel provides the facilities to operate in a parallel environment, such as message passing, prioritized process scheduling, and process creation and destruction, and memory management. The MGK hides the underlying parallel hardware structure by providing a virtual machine, the Multigraph Virtual Machine (MVM). The high level tools are constructed on top of this virtual machine. The user normally works with interactive graphical languages allowing him to concentrate on declaring a system model rather than program functions and control flow. These high level tools subsequently map the user's model of the required system into a parallel graph computation to be executed on the MVM.

The summer research focussed on the AEDC specific use of Multigraph Programming Environment, and the issues of an architecture for scaling the system to meet a real-world application of 24 full bandwidth channels.

II. OBJECTIVES OF THE RESEARCH EFFORT:

At the beginning of the summer, the basic research objectives were stated as:

- Demonstrate the existing system to potential users to gather feedback on processing capabilities needed/desired from the CADDMAS project.
- Develop methods and architectures for the efficient implementation of large-scale parallel signal processing systems. These should allow dynamic reconfiguration of processing algorithms and displays.
- Apply these concepts to the development of a demonstration system for the CADDMAS hardware.

Along with these objectives, a framework of the approach to be taken was sketched out as follows:

1. Fine-tune the demonstration prototype and software. Develop a demonstration and document the system in the form of a videotape to allow dissemination of the concepts to a wide variety of audiences. Present the system and gather feedback to drive the system specifications.
2. Develop architectures for efficient, cost effective implementation of large-scale parallel, interactive signal processing systems. The architectures should scale up to a large number of signals without a super-linear increase in the number of processors required.
3. Analyze the target application, a high-speed, high-bandwidth, many-channel real-time data analysis system for on-line analysis of turbine engine test data. Determine the requirements that the application places upon the software models and architectures.

This paper continues as follows:

1. Experiences with the Demonstration System are described. Feedback from the users is presented.
2. The Demonstration System Architecture is examined with respect to the scaling issues and architectural refinements are described to alleviate the problems.
3. Recommendations for further work are given.

III. DEMONSTRATION SYSTEM

One of the most challenging issues in this project is determining the user's needs. Since this effort represents a capability that is new to the engine testing industry, there is no body of users to query about desired features. The mind set of the analyst is predominantly oriented toward off-line processing and presentation of results. When asked about the desired features, a typical response is to ask what the system is capable of performing. The users usually underestimate what is feasible to in an online system.

Since the personnel involved in the development of the parallel signal processing system have little or no experience in engine testing, we must rely upon the users for input. To

assist the users in describing the ideal online processing system, we decided to develop a prototype that could demonstrate some of the capabilities of the system and would afford the user an opportunity for hands-on experimentation with the technology.

Early in the summer, we made minor adjustments to the software and refined the user interface to incorporate more of the underlying capabilities of the demonstration system. Briefings were prepared and a script for the methodical demonstration of the system's capabilities was constructed.

Two briefings were presented. The first briefing was presented to the AEDC potential users of the system. The second briefing audience was AEDC upper-level management and outside visitors from General Electric Evansville Division. The feedback was overwhelming positive. The engine analysts were amazed to see diagrams, which typically take 3 or more weeks to generate, appear online. The Campbell Diagram, which is the staple of the engine analyst's toolbox, had never been seen online before. The digital storage was of particular interest. This capability can streamline the offline analysis procedure. The most commonly requested feature that was not incorporated into the system was Engine Order Tracking, which plots the vibratory response of a blade at a multiple of RPM over a range of RPM. GE is preparing a formal response, which was not available at the time of this report.

A short, 5 minute, videotape was created to allow wide dissemination of the capability. We are optimistic that this will result in constructive feedback applied to the system requirements specification.

IV. ARCHITECTURAL ISSUES

The demonstration system has been developed as a result of the previous years of work. Refer to [3] and [5] for more details on the implementations of the processing elements (PE's) and their related software modules. The discussions presented here will be at the PE level and higher.

Figure 3 displays the Demonstration System Architecture. We can see the effect of signal flows on the block diagram. The system implements a pipeline style architecture to perform the standard, per channel operations on the signal. The FEP digitizes, filters, and monitors

the data. From there, the data is passed to the FREQUENCY processor for engineering unit conversion. Next, the data is passed to the IOPE, which stamps the systemwide time code, IRIG, on the data and splits the datastream into two paths, one goes to the Storage Processing Element, and the other branch goes to the rest of the processing pipeline. The T3 simply splits the data to the Crosscorrelation processor, and the Numeric Processing Element. The Crosscorrelation performs the signal processing crosscorrelation function in the time domain. The NPE computes a windowed FFT and the spectral magnitude. The NPE then locates the peaks in the spectral magnitude and passes this information out to the successive processors. The Campbell processor accumulates this peak information into a Campbell and Phase Campbell diagram buffer. All results are propagated at the appropriate times to the Graphics Processing Element for plotting, if necessary. The SPE can store all time domain data at the disk rate (800 KB/Sec) and can recover signals at a variety of speeds. No compression, other than byte packing occurs. The SPE implements a wrapping circular buffer. The MUX22 processors as data distribution nodes, ensuring that the data from any channel arrives at both of the graphics processors. The SIMPLE processor serves as a multiplexor to allow the PC to access and be accessed by three different transputers concurrently.

By examining this architecture, one can readily see that all of the functions in the pipeline can be incorporated into one single node, which we shall refer to hereafter as the **Integrated Sensor**. Independent of the implementation of its hardware, this single node can be applied to all input channels. The signal can then be conceptualized as not only a sensor data stream, but a predefined set of processed data sources. For instance, the Integrated Sensor can be the producer of time and frequency data, with no concern for the algorithms required to produce the data. This abstraction simplifies the higher level construction of data flow graphs.

One can also see, by examining the tree type nature of the Demo System Architecture, that communication saturation will occur as the number of signals is increased. Since all data from all sensors and the processing applied to the sensors is propagated upwards, the dataflows grow immense in short order. The current Demo Architecture has no provision for data flow control. Even data that is not currently being displayed is delivered to the graphics processor. While this functions well for a small number of channels, the method

does not scale. The solution to this problem is to provide a data metering capability on the integrated sensor. Only data that is needed is propagated onto the network. Data bandwidth is conserved for only the required data. No bandwidth is wasted.

V. INTEGRATED SENSOR SPECIFICATIONS

The major functions of the integrated sensor (IS) will be those processing operations that are required on a full-time, per-channel basis. The Integrated Sensor controls a single input datastream, which may correspond to one high speed signal, or a multiplexed set of low speed signals. The functions of the Integrated Sensor has the following requirements:

1. Signal Acquisition, with startup synchronization and a common acquisition clock.
2. Antialias Filters, to reduce the output of the integrated sensor to Nyquist rates.
3. Peak Detection, high accuracy, at the full system bandwidth.
4. Validation of Peaks with signal RMS values.
5. Thresholding of Peaks with configurable values for Alarm Generation.
6. Stamp data with Time, RPM, and other Relevant Transient Parameters.
7. Thresholding of Frequency Energy with configurable envelopes for Alarm Generation.
8. Fast Fourier Transformation to extract frequency domain information.
9. Data Storage and Retrieval of time domain signals.
10. Data Compression of time signals prior to storage. Compression *must* be lossless.
11. Data merge capabilities, to incorporate data streams from other Integrated Sensors.
12. A flexible data reporting facility, to reduce data bandwidth external to the Integrated Sensor.

The IS also has several inputs which can propagate results and receive information for combination with internal information. The following data packets are required:

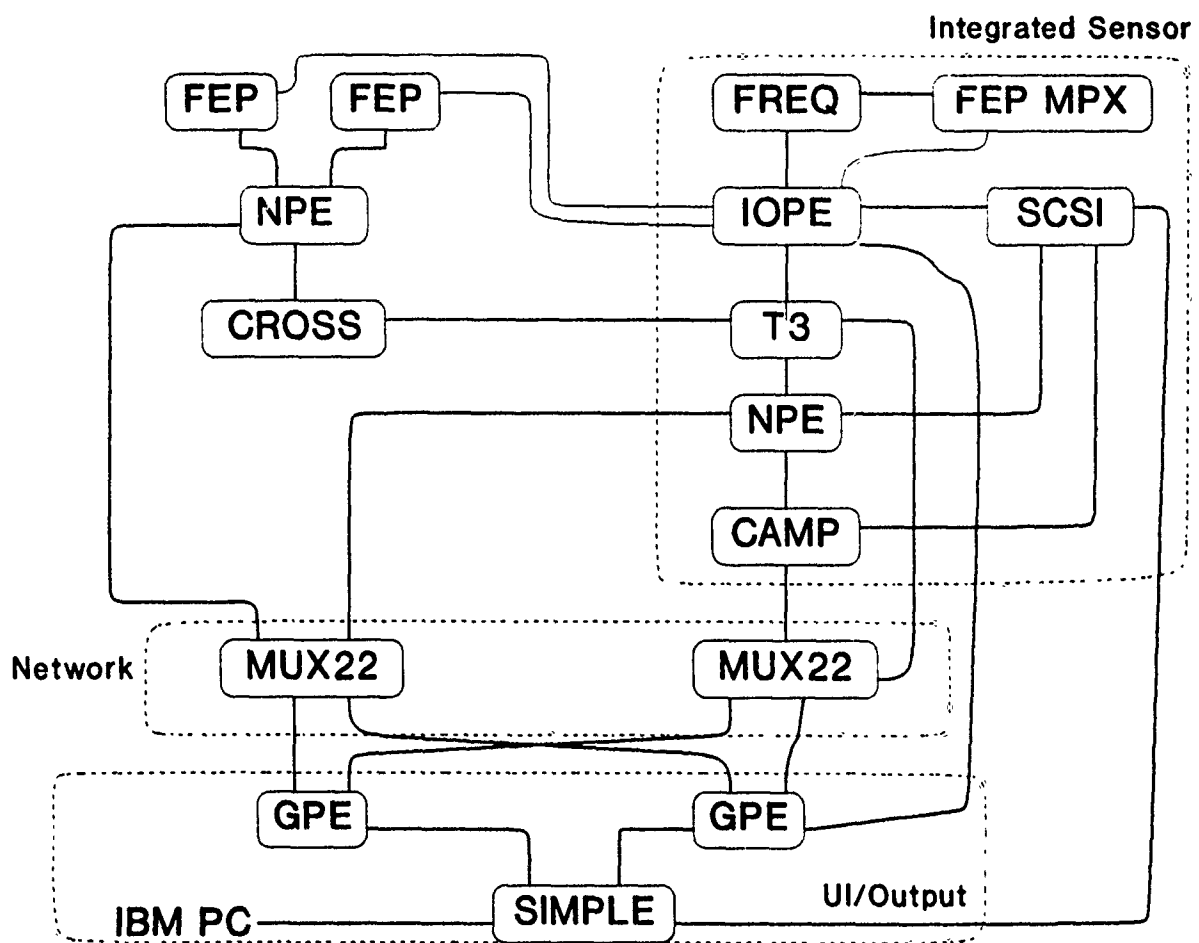


Figure 3: Demonstration System Architecture

1. **Raw Time**, the sampled input signal with calibrations applied, resulting in Engineering Unit values.
2. **Synchronized Time**, the calibrated input signal, with blocks synchronized to an external pulse. This results from the need to display the signals in reference to the once per revolution signal associated with rotating machinery.
3. **Correlated Time**, a combination of two synchronized time packets for later cross-correlation. Data synchronization is achieved by a common input sampling clock, common data block sizes, and a synchronized starting time. Data blocks remain synchronized.
4. **Incomplete Correlated Time**, a time packet to be used to create a Correlated Time packet.
5. **Complex Frequency**, the result of a Fast Fourier Transform on the input signal.
6. **Frequency Energy**, the real magnitude of the FFT output.
7. **Frequency Energy Peak Hold**, the peak-hold of the magnitude of the FFT output over time.
8. **Frequency Tracking**, the magnitude of the frequency spectrum at a multiple of engine RPM.
9. **Frequency Energy Peaks**, the top N peaks of the frequency magnitude spectrum. N is a configurable parameter. The peaks consist of a frequency value and the energy of the signal at that frequency. These values will be the result of a parabolic interpolation to improve accuracy.
10. **Phase Frequency Peaks**, the complex FFT values at the peak magnitude frequencies of a primary signal, along with the complex values of a secondary signal at those peak magnitude frequencies. This packet is used to generate Phase Campbell Diagrams. The secondary signal emanates from another integrated sensor.
11. **Incomplete Phase Frequency Peaks**, the complex FFT values at the peak magnitude frequencies of a primary signal are sent for combination with the secondary signal's complex frequency values.

12. **Absolute Value Alarms** , reported when the signal value exceeds a configurable threshold.
13. **Frequency Envelope Alarms** , reported when the signal energy exceeds a configurable threshold. The threshold envelope is specified for each frequency bin in the spectrum.

All of the parameter packets are calibrated into engineering units. Each data stream associated with a packet has several parameters. To reduce the bandwidth requirements, the data reporting rate is specified. This is in the form of a modulo factor, selecting one in M data blocks. Using this rate, the external bandwidth requirements are those of the consumer. For slow consumers, such as graphics devices, a tenfold or greater reduction in bandwidth occurs. The data stream parameters also specify the destination of the data, so the information can be routed directly to the consumer.

VI. SOFTWARE ENVIRONMENT:

It becomes apparent from the application area that several unique demands are placed on the system. There are an extremely large number of configurations of processing algorithms possible for the system at any one time. With thousands of inputs, and multiple algorithms to execute, there is a combinatorial explosion of possible configurations. Handling the large number of possible configurations can be an enormous task if all of the combinations are implemented statically. This problem can begin to be approached if the system is capable of interactive algorithm reconfiguration. This capability, however, imposes significant demands on the system. The Probes system is being designed to provide a framework wherein these problems associated with these requirements can be solved in an organized manner.

The conceptual implementation of these facilities have been refined over the course of the summer. The previously described Integrated Sensor model, provides a clean interface to the data sources. The data streams can be dealt with in a high level manner. The facilities for opening a synchronized data stream, especially for dual channel operations, greatly reduces the real-time requirements imposed upon the general purpose segments of

~ the system.

Signal processing systems are to be constructed by connecting the various types of signal processing modules in an arbitrary structure. The Multigraph Kernel provides these services [2]. The process occurs as follows:

1. The user interface requests a specific type of online plot, and specifies the resolution and temporal update rate.
2. The Builder recalls the necessary processing module structure from its library of pre-built structures.
3. An evaluation of the requested resources vs. those available assesses the feasibility of adding this structure to the global processing system.
4. If feasible, the Probes Builder will create the necessary computational actors and associated datanodes. Part of every structure will be an interface to the integrated sensor.
5. The interface to the Integrated Sensor will use the Streams facility provided with the Multigraph environment. A real-time actor will monitor the datastream in the form of a Streams input port. This interface will convert the datastream into a standard MGK datanode.
6. The builder will open the necessary datastream on the Integrated Sensor and initiate data flow.
7. The Builder will send the appropriate commands to the GNUPlot graphics interface to display the data stream.
8. Finally, all of the processing structures will be enabled and data flow will commence.

The probe is designed in an object-oriented manner. The probe object will contain information about processing algorithms, supported data types, required resources, processor assignment hints, initialization requirements, etc. The object's information can be used for automated processor placement. It will be the task of the Builder to use this information, in combination to the current state of the processing hardware, to perform an acceptable assignment of processing algorithms to processors

The user interface allows the user to install and connect signals, processing operations, and displays in the signal processing system. Given the feedback from the Demonstration System, it appears that a simple, text-menu based interface will be more efficient than a graphical signal-flow depiction language. The details of implementation are not important to the future users of the system. In fact, a graphical signal representation is likely to confuse the users, most of which have little formal training in signal processing and computer systems integration.

The Probes Builder is currently under development and will continue under the AFOSR Mini-Grant funds. The final version will be a C++ implementation due to the object-oriented nature of the Probes and the user interface required.

VII. ADDITIONAL SUMMER TASKS:

In addition to the work on the major focus described above, the following peripheral work was completed:

1. Interfacing with DDAL personnel to gather additional specifications for the CAD-DMAS system.
2. Consulting on the specifications and development approach for the 24 channel ATF system.
3. Harvard Graphics figures for the Demos were generated for display on the Graphics Processing Element.
4. A briefing was prepared and presented to the Advanced Supercomputer Evaluation Project Committee (ASEP) on the requirements associated with the Dynamic Data application area.
5. Installation/Upgrade of the Unix systems software on the IBM RISC/6000 workstation.
6. A copy of the public domain GNU C/C++ 386 compiler was distributed to several

groups in Svt. Technology. The flexible access to large memory and the speed increase will allow them to fully exploit their Intel 386 processor PC's for large problems.

7. The 1990 generation Front End Processor was validated for antialiasing and processor synchronization. A single A/D converter clock should be distributed to all boards to ensure that the channels remain in sync.
8. A presentation by the NCUBE corporation was attended. An evaluation of the architecture and systems capabilities with respect to the Dynamic Data application was prepared.
9. Assistance was given to finalize the layout for the Graphics Processing Element.

VIII. RECOMMENDATIONS:

In summary, engine testing applications require extremely demanding computational resources. The structure of these applications is well suited to parallel architectures. The high data rates and the need for dynamic reconfiguration of algorithms complicates matters beyond the usual issues in parallel program development. The combination of the Integrated Sensor architecture and the reconfigurable probes described here works well in this environment.

Further work needs to address the following issues:

- The Integrated Sensor implementation based on the currently available hardware has been started and should be completed. This will flesh out all of the minor details associated with this approach. Preliminary work can be accomplished on the workstations. Final implementation will require the CADDMAS hardware.
- The tasks performed by the Integrated Sensor are performed very efficiently by processors optimized for signal processing. The Integrated sensor should be implemented in DSP technology to optimize performance and cost. This implementation will rely heavily on the results of the implementation on the current hardware.

- Complete the definition and implement the Probes Builder. The builder can be implemented/debugged in a software environment on the workstations.
- Modify GNUPLOT to support the MGK dataflow interface, for use in a pure MGK environment. The GNUPlot should also be adapted to support 3-D graphics for the display of Campbell Diagrams and other multidimensional data.
- The installation of the CADDMAS system in the DDAL needs to be supported. Invaluable feedback will be received from this effort.
- Load balancing, performance monitoring, and debugging tools are needed to optimize the usage of the CADDMAS processors. This research needs to be integrated into the framework of the Probes environment. Information of the problem is vital to making decisions about processor placement and communications balancing.
- Issues of dynamic processor/sensor replacement and management should be investigated due to the large number of sensor failures in real aero-mechanical tests. Also, fault tolerance and hardware diagnostics issues must be addressed when dealing with a processing system of this magnitude.
- Continue evaluation of new and evolving system processors and architectures. Contenders in this area are the next generation Inmos processor (the T9000), the Intel I860XP, and the Intel IWARP

Acknowledgements

We wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Research & Development Labs must also be commended for their concern and help to us in all administrative and directional aspects of this program. Our experience was rewarding and enriching because of many different influences. Lt. Greg Nordstrom provided us with support, encouragement, and a truly enjoyable working atmosphere. The help Tom Tibbals was invaluable in overcoming many technical roadblocks. Jason Scott of the High School program did a terrific job. Finally, we would like to thank Major Hart for all his help in insuring that our experiences at AEDC would be memorable.

References

- [1] Biegl, C.: "Design and Implementation of an Execution Environment for Knowledge Based Systems" Ph.D. thesis, Department of Electrical Engineering, Vanderbilt University, Nashville, TN., Dec., 1988.
- [2] Abbott, B., Biegl, C., Sztipanovits, J.: "Multigraph for the Transputer".
Proceedings of the Third Conference of The North American Transputer Users Group.
Santa Clara, Calif., April 1990.
- [3] Abbott, B., Bapty, T.: "High-Speed Parallel Signal Processing",
FINAL REPORT. 1990 USAF-UES SUMMER FACULTY RESEARCH PROGRAM.
Conducted by the Universal Energy Systems, Inc. Contract No: F49620-88-C-0053.
Aug., 8, 1990.
- [4] Bailey, D., Barton, J.: "The NAS Kernel Benchmark Program", National Aeronautics and Space Administration, Technical Memorandum. Aug. 1985.
- [5] Nordstrom, G. AEDC/DOTR: "I/O Interfacing for the Transputer"
Proceedings of the 23 Southeastern Symp. on System Theory.
Columbia, South Carolina. March, 1991.
- [6] MOTOROLA: DSP56000 Digital Signal Processor User's Manual, 1986.
- [7] INMOS Limited: Transputer Reference Manual, 1989.
- [8] Karsai, G.: "Hierarchical Description Language (HDL) User's Manual" Dept. of Electrical Engineering, Vanderbilt University, Technical Report #87-004, 1987.
- [9] Texas Instruments Incorporated: TMS34010 User's Guide.
- [10] Zoran Corporation: ZR34345 32-Bit IEEE Floating-Point Vector Signal Processor User's Manual.

WAKE AND PROJECTILE VELOCITY ESTIMATION

D. Mitchell Wilkes & Georges Badih Aboutanos

Vanderbilt University

Department of Electrical Engineering

Nashville, TN 37235

1 Abstract

Digital signal processing techniques were used to estimate the velocity of a projectile and its wake. The observed data was collected by one or multiple doppler radars in an underground ballistic range.

A classical FFT-based spectral estimation approach was used to identify the frequency characteristics of the data. These frequencies were later translated into velocities leading to accurate velocity profiles of the projectile and of the wake.

A C program was also developed to implement the spectral estimation approach and to provide an automated analysis of the data with extensive graphical display of the results. These graphical presentations include velocity profiles, 3-D surface plots, and contour plots.

2 Introduction

The main goal of this research was to develop an automated analysis technique for doppler radar data from a ballistic range. The doppler data is collected on a projectile followed by its wake, then it is automatically analysed to provide a graphical and numerical description of the velocity behavior of the projectile as well as the wake, which is more difficult to analyze.

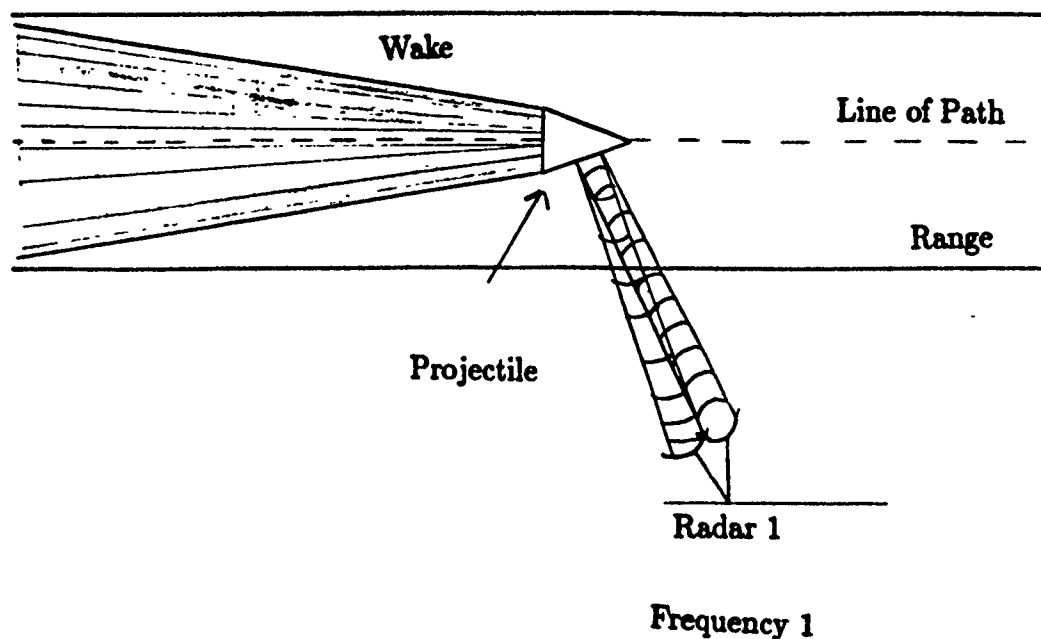


Figure 1: Projectile velocity measurement.

2.1 Physical Description of the Range

The measurements are taken in a long underground range, usually evacuated to a low pressure. A typical projectile is about 10 cm long and 2 cm in diameter is fired down range by a two-stage compressed light-gas gun at a speed of approximately 20,000 ft/s. As the projectile travels down the range a few doppler radars, usually three, each at a different frequency (8.6, 17 and 35 GHz), measure the velocity of the projectile and the wake following the projectile. These radar antennas are usually located about 45° from the line of path of the projectile. (See Figure 1)

2.2 Description of the Measurement System

This section describes the radar measurement system and the way the doppler frequency is measured and converted to velocity. The radar is transmitting at some frequency f_r . This is a quasi-monostatic radar system where the transmitter antenna and the receiver antenna are co-located. When the radar wave hits the

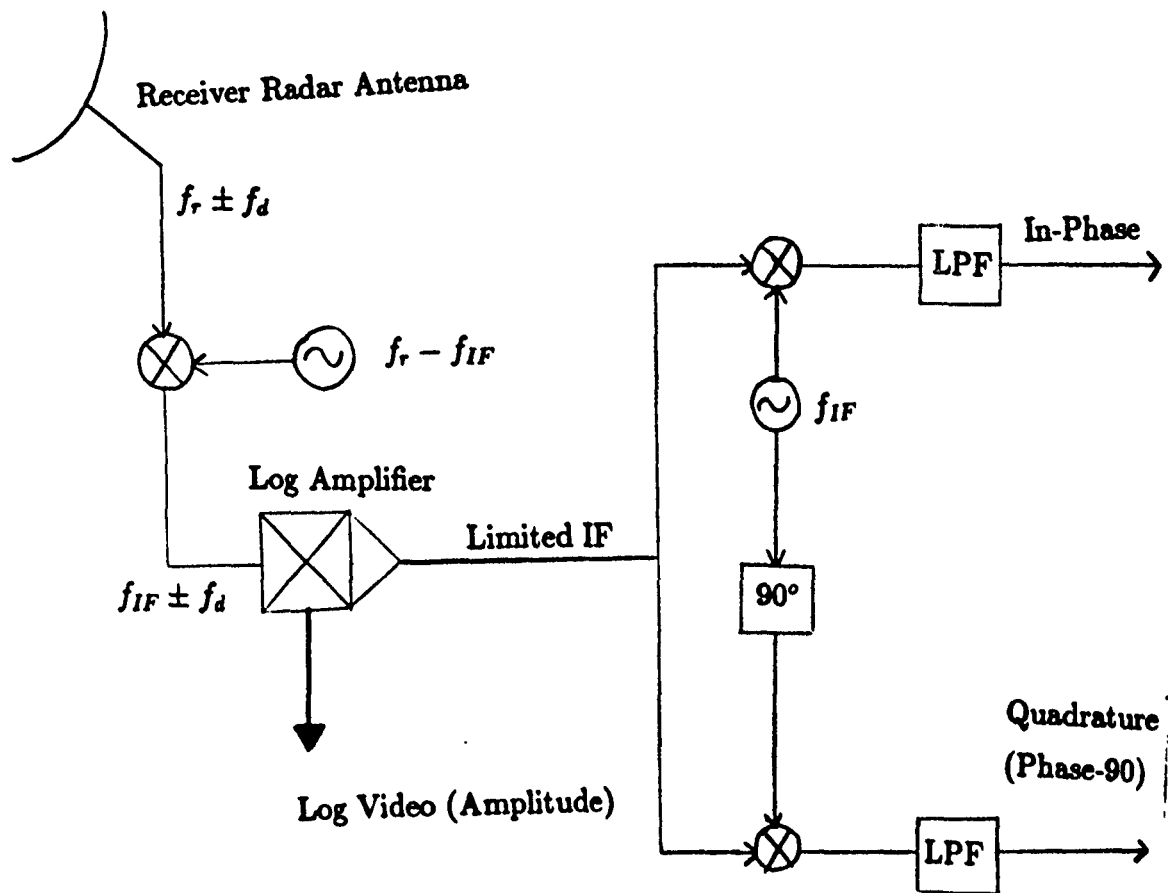


Figure 2: Radar measurement system.

moving projectile or its wake it reflects back to the radar receiver at the frequency f_r shifted up by a doppler frequency f_d .

The received frequency $f_r + f_d$ is down converted to an intermediate frequency $f_{IF} + f_d$. A Log amplifier is used to collect amplitude data, and the limited IF data is separated into an In-phase channel and a Quadrature channel (which is phase shifted by 90° relative to the In-phase channel). The data is then downconverted to DC by multiplying it again with quadrature sinusoids at f_{IF} , and finally each channel is passed through a low pass filter leading to a doppler frequency related to the speed of the projectile and its wake at the In-phase and Quadrature channels (See Figures 2 and 3). LPF is the low pass filter, f_{IF} is the intermediate frequency, f_r is the radar frequency and f_d is the doppler frequency.

The doppler frequency is converted to velocity using the following equations [1]:

$$f_d = \frac{2v \cos(\theta) f_r}{c} \quad (1)$$

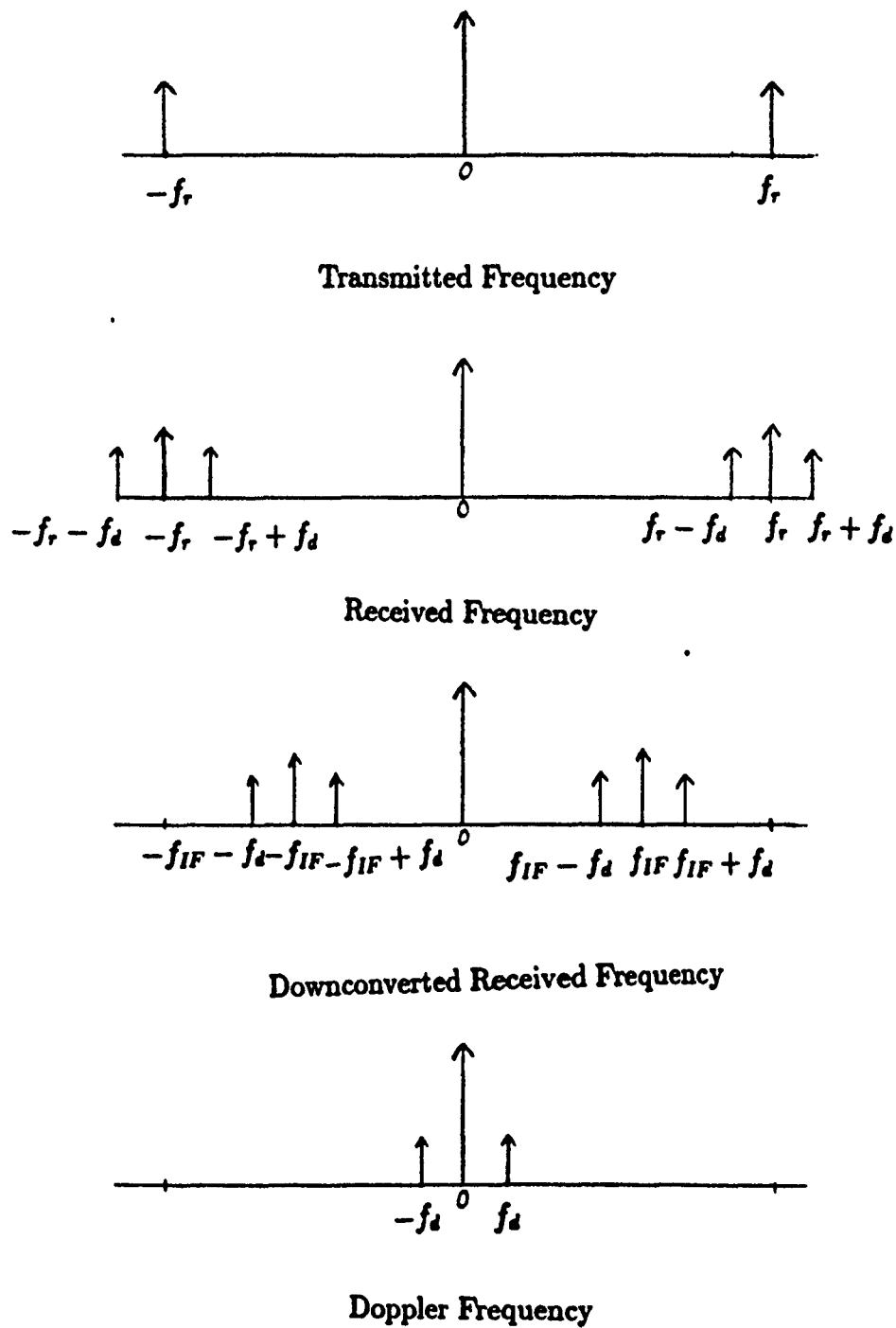


Figure 3: Radar frequency measurement.

$$v = \frac{cf_d}{2 \cos(\theta)f_r} \quad (2)$$

where v is the velocity of the target, θ is the angle between the line of the path of the projectile and the radar line of sight (usually 45°), c is the speed of light, $3 * 10^8 m/s$, f_r is the radar frequency, and f_d is the doppler frequency.

Substituting f_d into Equation (2) will provide the velocity of the projectile or of the wake, v . The velocity and the doppler frequency are proportional, which means the larger the doppler shift the faster the projectile is moving and the smaller the doppler shift the slower the projectile is moving.

2.3 Description of the Projectile and the Wake Results

The raw data is nominally taken at a sampling rate of 200 ns per point. This data typically consists of four sections. The first section is random noise at low amplitude, the second section is the return from the projectile, which is the first object of interest seen by the radar. This section is characterized by a large constant magnitude and a high doppler frequency. The third section is the return off the wake and it has basically the same constant magnitude behavior as the projectile, but the frequency is lower because the wake moves slower than the projectile. Finally the fourth section of the data is random noise again at low amplitude. An example of such data for 6000 points and a 200 ns sampling period is shown in Figure 4.

The projectile and the wake can be distinguished easily from the background noise because they have a higher amplitude. After converting the raw data to velocity we observe a null area immediately behind the projectile that is due to the recombination of gases. This feature can be used to separate the wake from the projectile so that the velocity behavior of each (i.e., projectile and wake) can be observed separately. The final result will be displayed in terms of percentage of the model speed versus distance from the beginning of the projectile measured in projectile body diameters. A sample of such a result is shown in Figure 5.

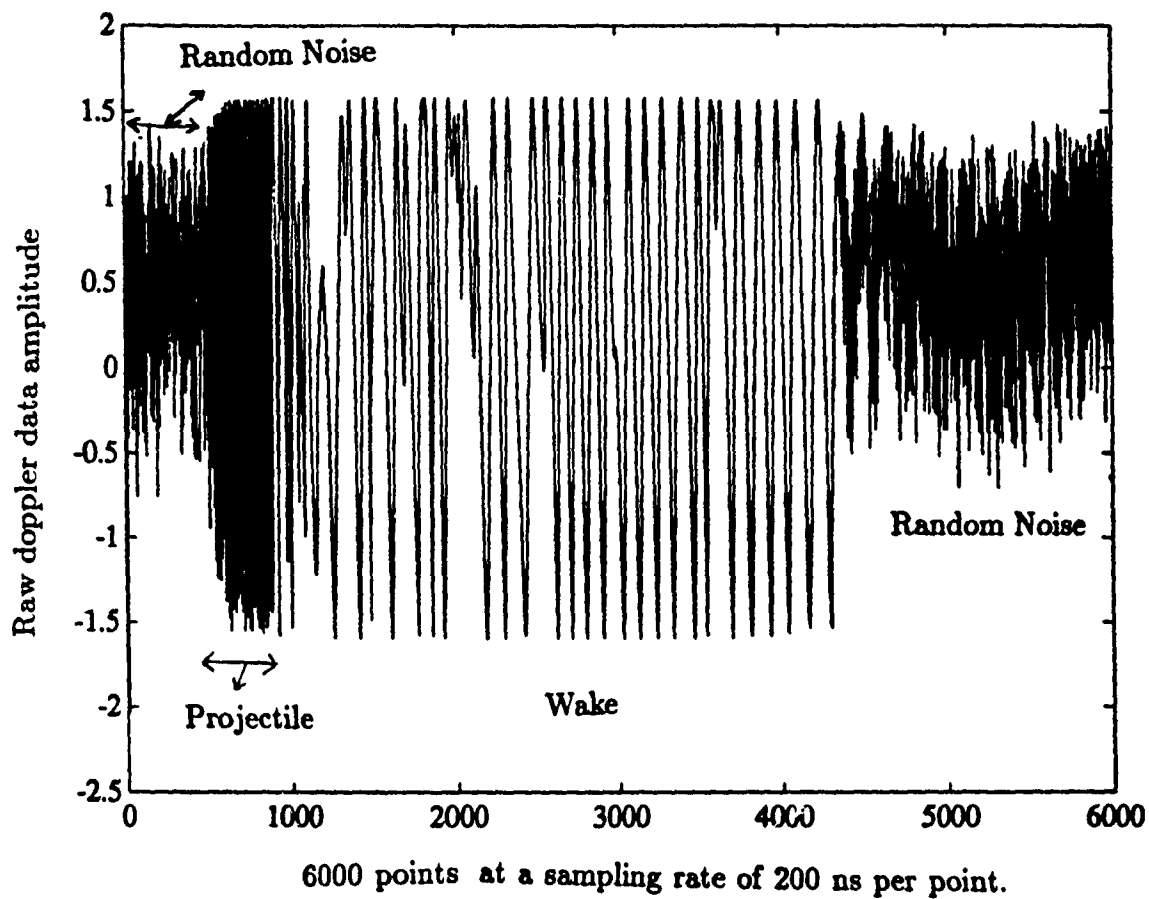


Figure 4: Raw doppler data.

PROJECTILE+WAKE (17.0 GHz)

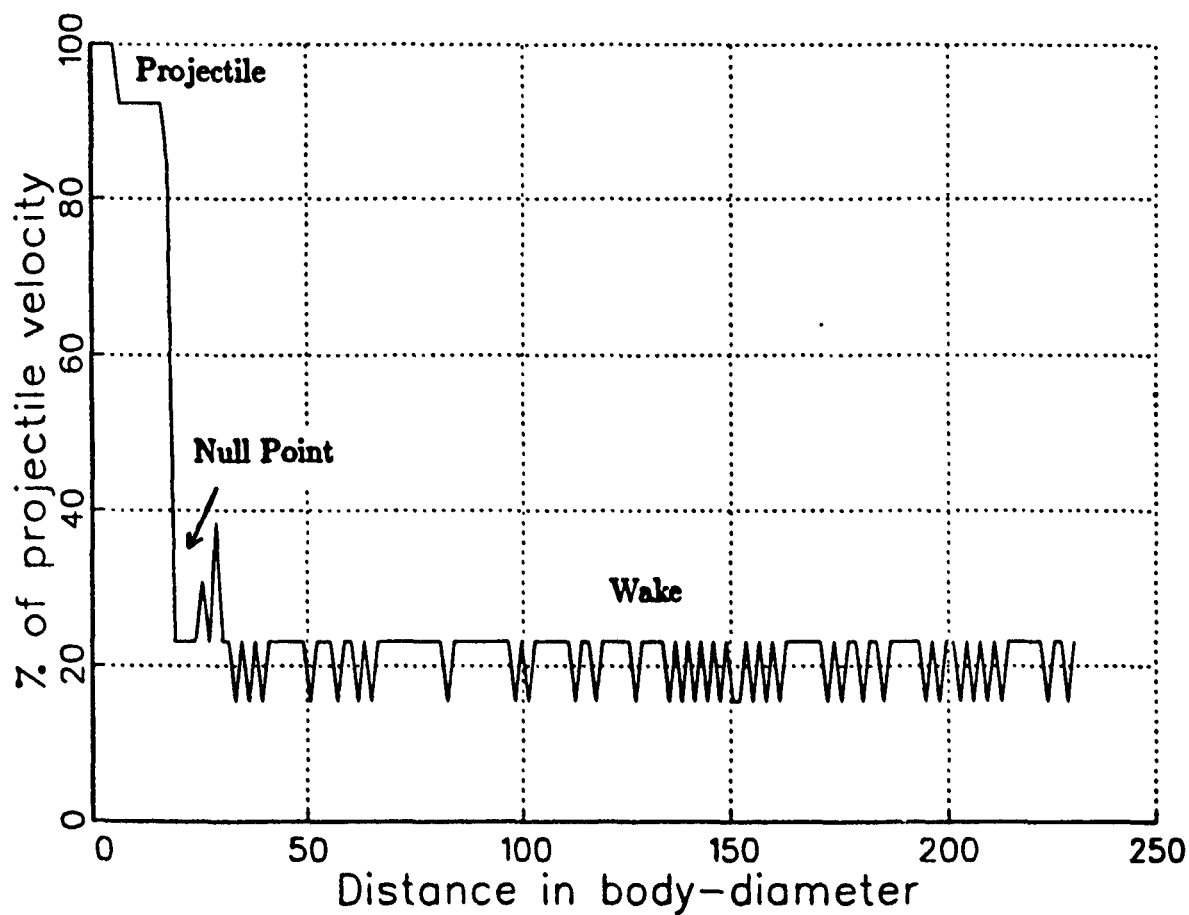


Figure 5: Extracted projectile and wake velocity profile.

3 Problem Discussion

3.1 Statement of the Problem

Provided with some doppler data collected on a projectile followed by its wake, our goal was to use digital signal processing techniques to estimate the velocity of the projectile and the velocity of the wake for possibly multiple radar frequencies such as 8.6, 17 and 35 GHz. The analysis of the data and the estimation technique should be fully automated providing a projectile and wake velocity profile, and a wake velocity profile alone. New visualization techniques were also proposed to provide a clearer understanding of the projectile and wake velocity behaviors.

3.2 Basic Approach

The basic approach of this effort was as follows. The doppler data is broken into many possibly overlapping windows. Spectral estimation is performed separately on each window. We find the location of the tallest spectral peak in each window and use it as our estimate of the doppler frequency for that time window. The peaks from each time window are collected and represent the doppler frequency behavior as a function of time. These doppler frequencies are then translated directly into velocities via Equation (2), and displayed in different graphical formats describing the velocity behavior of the data. Analysis is performed to extract the velocity profile of the projectile and the wake and to distinguish them from the background noise. This basic approach will now be described in greater detail.

3.3 Details of Spectral Estimation

Spectral estimation is used to identify the frequency characteristics of the data, which can be translated directly into velocity. A classical FFT-based spectral estimation approach was chosen due to its well behaved characteristics [2]. Although this approach may have difficulty resolving closely spaced sinusoids, this limitation does not cause a problem for a basic analysis of the data.

The first task is to break the observed sequence, $x(n)$, into many possibly over-

lapping L point long windows, $x_m(n)$, given by

$$x_m(n) = \begin{cases} x(mM + n) & n = 0, \dots, L-1 \\ 0 & \text{else} \end{cases} \quad (3)$$

where M is the distance in samples between the beginnings of successive windows and mM is the location of the beginning of the m^{th} window. Before applying FFT-based spectral estimation on each window, we applied window functions to reduce the ripples in the estimated spectrum due to sidelobe effects from using finite length data windows. Some of the windows that were used include the Hamming, Hanning, Blackman, and Rectangular windows [2]. These windows also provide a degree of spectral smoothing that tends to suppress false spectral peaks. Each L point long window is multiplied by the chosen L point long window function $w(n)$ to form

$$y_m(n) = x_m(n)w(n) \quad (4)$$

The squared magnitude of the FFT of $y_m(n)$ is used as the estimate of the power spectrum of the m^{th} window, given by

$$S_m(k) = |Y_m(k)|^2 = \sum_{n=0}^{N-1} y_m(n) e^{-j\frac{2\pi}{N}kn} \quad (5)$$

N is the length of the FFT and should be a power of 2 in order to be able to use the common radix-2 fast algorithms [2]. We also require $N \geq L$, and if $N > L$, zeros should be appended to $y_m(n)$ until its length is N . The frequency bin number k corresponds to a discrete-time frequency of $\frac{2\pi}{N}k$.

We further form a spectral matrix with columns consisting of the successive $S_m(k)$'s computed from the time windows. The major advantage here is that this allows us to observe all the frequencies in the data and not just the peak doppler frequencies. This provides a good 3-D view of the evolution of the doppler frequencies (i.e., velocities) over time (see Figures 6 and 7). False peaks are also suppressed by averaging the columns in the matrix three at a time to achieve some degree of temporal smoothing in the frequency domain.

PROJECTILE+WAKE (17.0 GHz)

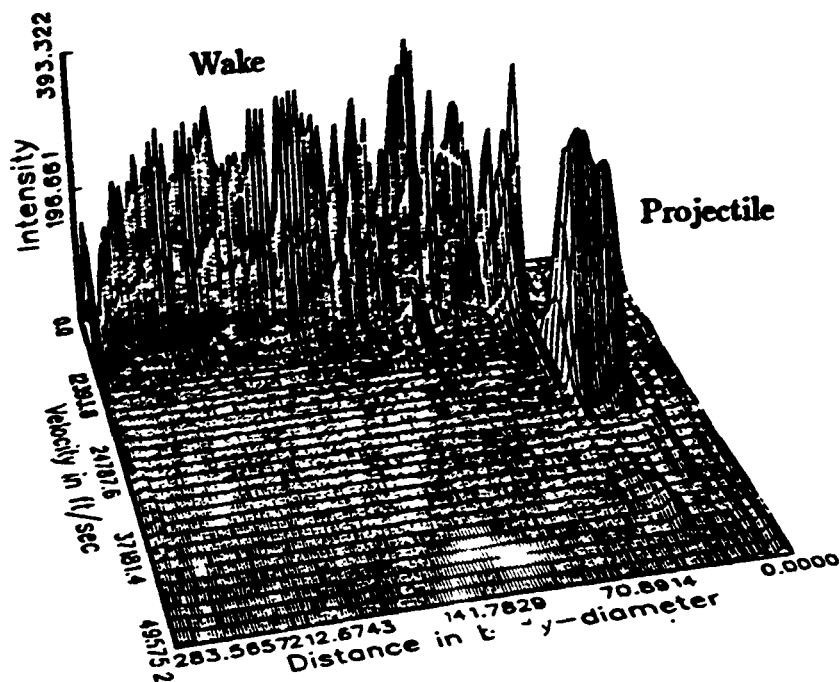


Figure 6: 3-D plot of the evolution of doppler frequency with time.

PROJECTILE+WAKE (17.0 GHz)

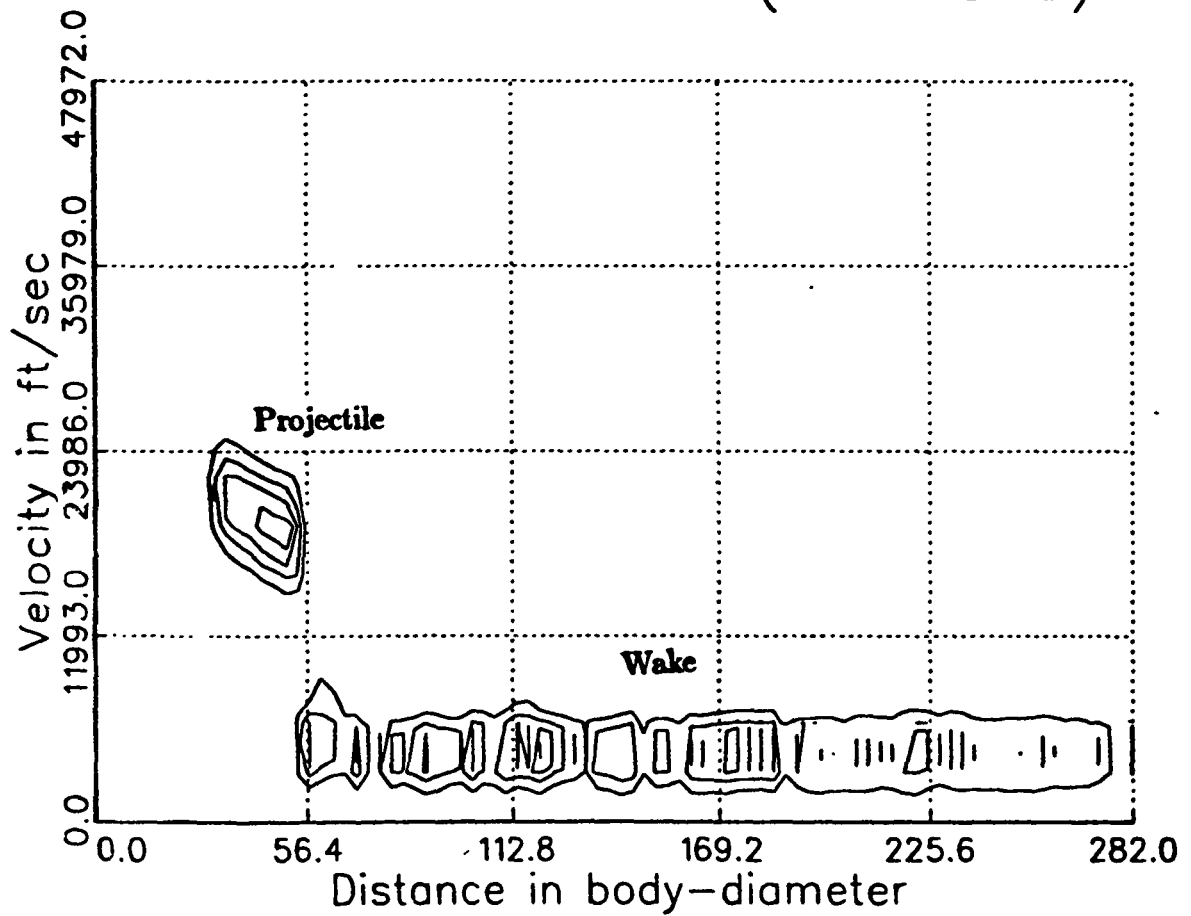


Figure 7: Contour plot of the evolution of doppler frequency with time.

3.4 Automatic Extraction of Projectile and Wake

In section 2.4 we discussed the characteristics of the wake and the projectile. In this section we show how these characteristics are used to separate the projectile and the wake from the noise and how the projectile and the wake are automatically separately from each other.

When we observe the raw data we notice that the projectile and the wake have a higher and more constant amplitude than the random noise in the beginning and at the end of the data. Therefore, we locate the maximum point, S_{max} , in the spectral matrix (Figure 6) and assume it is located on the projectile or the wake (this has always been observed to be the case). Then, we start comparing the maximum peak amplitudes of the $S_m(k)$ power spectrum estimate from each window to a threshold (for example $S_{max}/3$ provided good results). The location where the values cross this threshold determines the beginning of the projectile. All the data before that point is considered to be noise.

A similar approach is used to locate the end of the wake. However, we do not compare each peak of the $S_m(k)$ individually, but rather we create a test window and compare the maximum value of that window to a threshold (for example $S_{max}/4$ provides good results). Such a window approach is necessary due to the nature of the wake data which may drop below the threshold for a short time. A window will detect whether the data is below the threshold for a longer period of time. Using the two points identified in this way we extract the data from between and consider it to be the data for the projectile plus the wake (See Figure 8).

As was mentioned earlier, after we convert the raw data to velocity we observe a null area immediately behind the projectile. This characteristic plus the fact that the wake should be traveling slower than the projectile provide a way to extract the wake data using a frequency threshold rather than an amplitude threshold. The threshold used here or when the data is being separated from the noise is very dependent on the nature of the data (see Figure 9, the extraction of the wake).

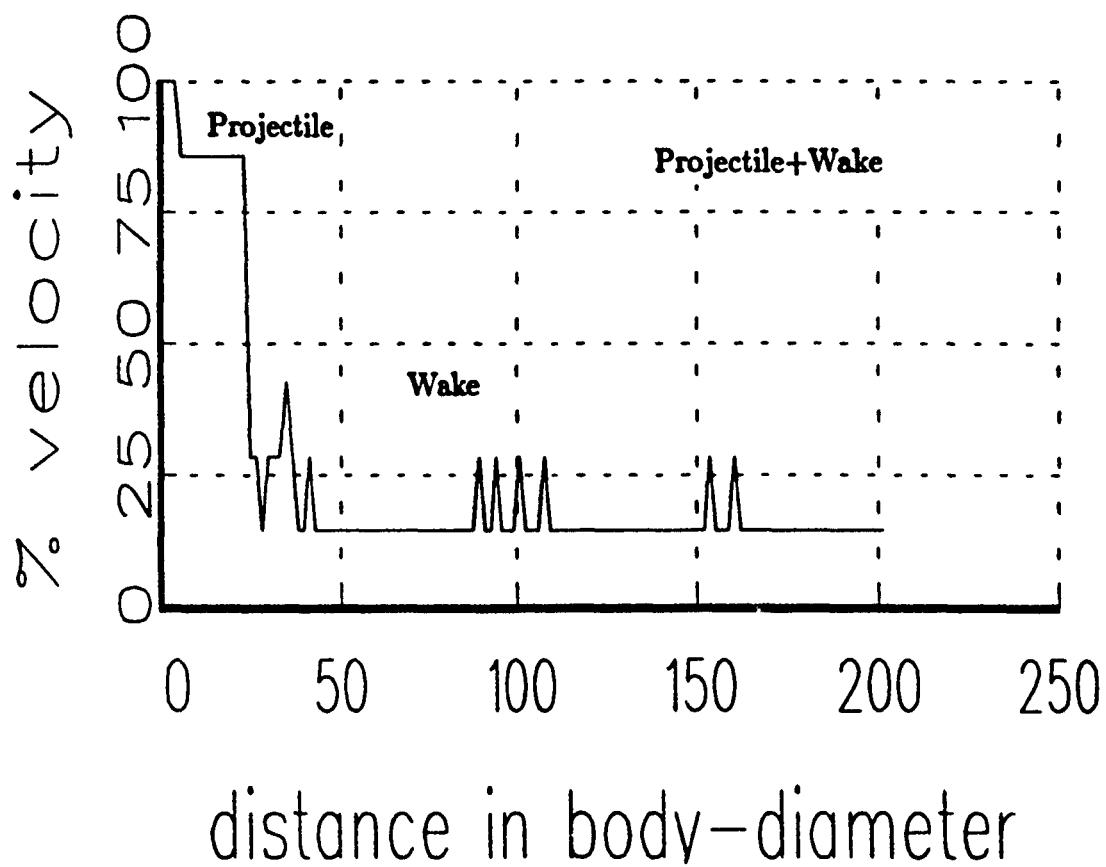


Figure 8: Extracted projectile plus wake velocity.

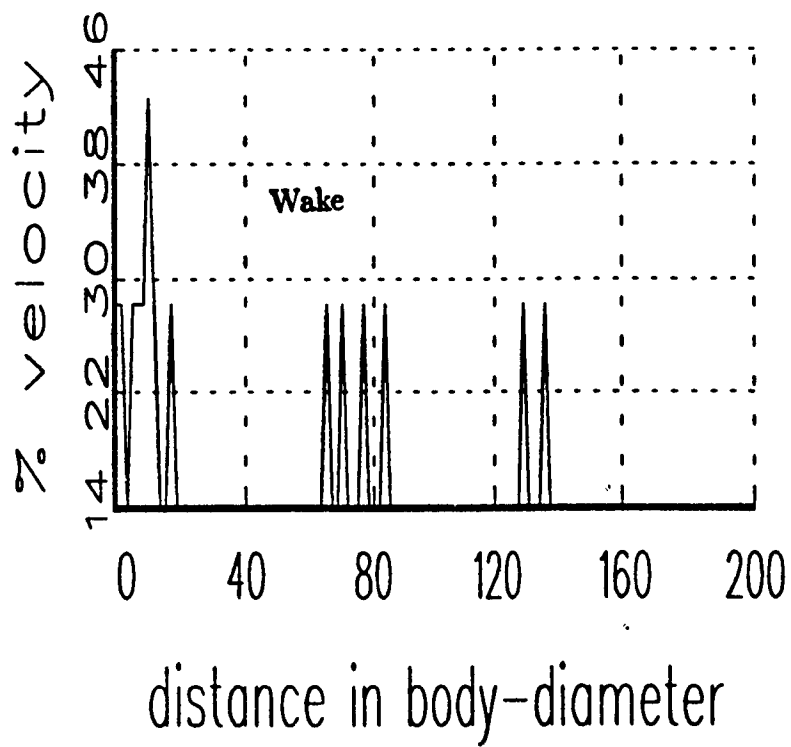


Figure 9: Extracted wake velocity.

4 Results

We developed a C program to implement this spectral estimation approach. This program was designed to be user-friendly, and provides an automatic and complete graphical analysis of the velocity behavior.

The user is prompted for a few inputs such as the name of the file containing the measured data, the projectile body diameter, the sampling period, the smoothing window preferred, the radar position and frequency and some other necessary parameters. The C program will then automatically extract the projectile and wake from the data and provide the user with detailed graphical presentations of the velocity behavior of the projectile and the wake. Two of these representations are 2-dimensional plots of velocity versus time. As mentioned earlier the velocities are normalized to the maximum velocity of the projectile and the time is converted to distance in terms of projectile body diameter. The other two graphical representations consist of the 3-dimensional plot of velocity, time, and intensity, and the contour plot. An example of these four plots is given in Figure 10.

These graphical presentations provide the user with revealing views of the data behavior. The program is also capable of comparing the results of a test done using three radars each with a different frequency. The output will be a set of eight plots, four for the projectile plus wake and four for the wake only. This capability allows the comparison of velocity behavior for different frequencies when performed on the same test. This technique can also be used to compare the data for the same frequency but from the In-phase, Quadrature and complex (In-phase and Quadrature data combined) outputs (See Figure 11).

5 Conclusion

The proposed approach can be used for automated analysis of doppler radar data. As mentioned in section 4, a relatively small number of initial parameters are needed to provide an automated graphical analysis of the velocity of the projectile and its wake. The proposed means of visualizing the results through 3-D surface plots

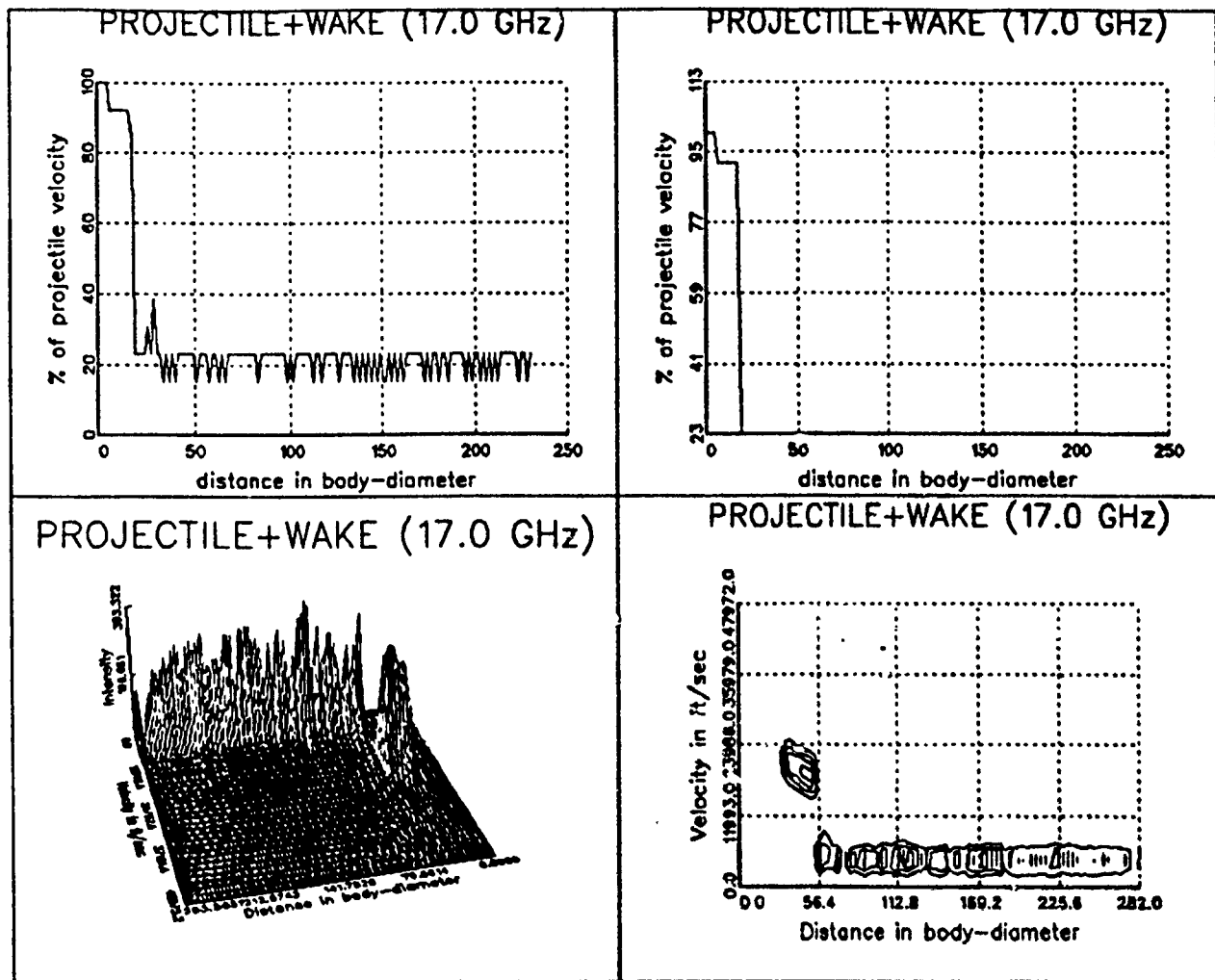


Figure 10: Graphical presentation of results.

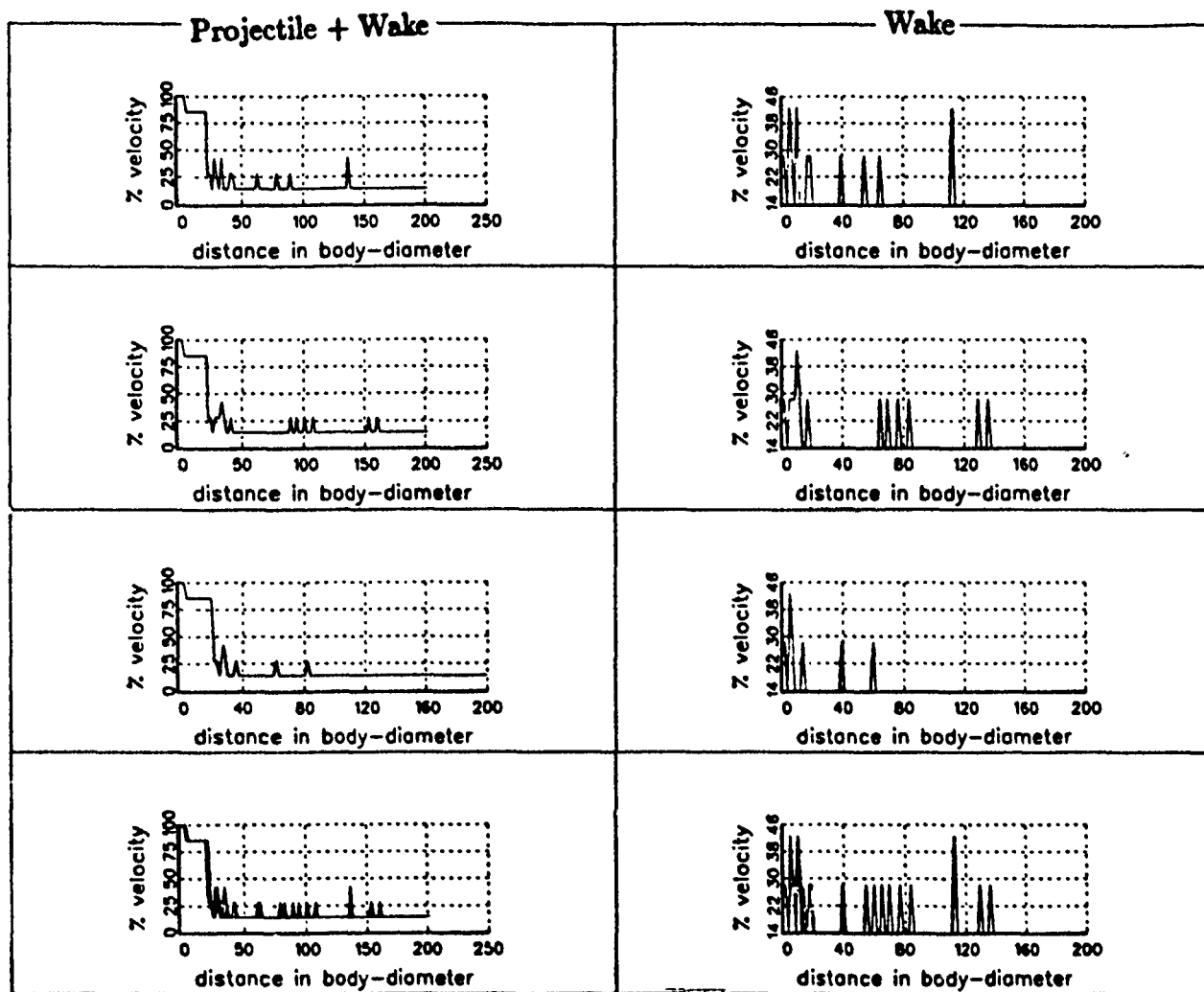


Figure 11: Spectral estimation comparison for 3 sets of data.

and contour plots provide a more complete understanding of the data. Future work may include the use of higher resolution and/or nonstationary spectral analysis techniques such as auto-regressive modeling and time-frequency distribution techniques such as the Wigner-Ville transform that hold promise of increasing the accuracy of the frequency estimates.

Other applications should be studied such as in-barrel doppler data or 3-D imaging of the wake using multiple radars. The latter application is a Phase Interferometer that may be used to characterize qualities of the wake such as its diameter and the density of free electrons.

References

- [1] *Radar Handbook, 2nd Ed.* ed. by M.I. Skolnik, McGraw-Hill: New York, 1990.
- [2] A.V. Oppenheim and R.W. Schaffer, *Digital Signal Processing*. Prentice-Hall: Englewood Cliffs, 1975.

AN EXTENDED KALMAN FILTER OBSERVER FOR AN ALTITUDE TEST CELL

**D. Mitchell Wilkes and W. Brian Ball
Vanderbilt University
Department of Electrical Engineering
Nashville, TN 37235**

Abstract:

New control techniques are needed to keep pace with the development of technology in altitude test cells. A recently proposed technique for these controls is a model-following controller using an inverse-process model, but there may be room for improvement in the accuracy of the inverse-process model. During this effort, an extended Kalman filter was developed to reduce errors in the inverse-process model.

Introduction:

As simulated altitude tests are required to become more necessary and economically desirable, the systems required to control these test cells are also having to become more accurate.[1] The technology and techniques now available allow more realistic modelling and estimation of real-world conditions in test cells. Of special interest is the performance testing of engines under transient conditions. With costs of flight testing becoming much more expensive, economical alternatives, especially ground testing, become very important considerations.

The most common controller presently is a PID controller, sometimes gain scheduling is added to the PID controller to increase its flexibility [1]. A recent paper by Chaney [1] proposed an improved technique for controlling tests in these cells. The paper found a model following controller to be the most flexible and accurate form of control in this situation. The control consisted of a second order linear model generating smooth trajectories for each setpoint, using measurements or estimates of process states and parameters in an inverse process model to establish the control inputs to the plant. The controller compensates for nonlinear plant dynamics by linearizing and decoupling the system. Perfect modelling is the goal, but

tracking errors in the system develop from unmodelled dynamics or parameter errors in the feedforward controller. [1]

One source of error results from errors in the observation and estimation of the necessary parameters for the inverse process model. These errors can be reduced by employing an extended Kalman filter observer to estimate these parameters from the observed data. Such a Kalman filter was developed during this effort.

The next section provides a general discussion of the problem at hand. The sections following provide a general procedure for applying an extended Kalman filter to this problem and a specific procedure for developing the Kalman filter is given. A results section then shows the performance of the Kalman filter against simulated data.

Problem Discussion:

Our task was to design an extended Kalman filter to observe the important parameters of a altitude test cell. The Kalman filter observer will become part of a controller for the test cell. [1]

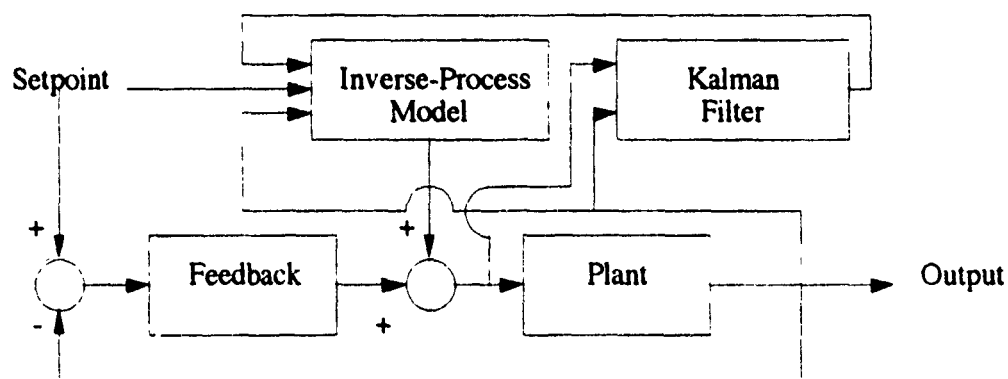


Figure 1 - Model Following Controller with Kalman Filter

The controller shown in Figure 1 is a model-following controller that includes a feedforward inverse-process model to translate the setpoint into

the proper inputs to the test cell. The Kalman filter is necessary to provide high quality estimates of the important parameters so that the feedforward inverse-process model will perform accurately. The test cell under consideration is shown in Figure 2.

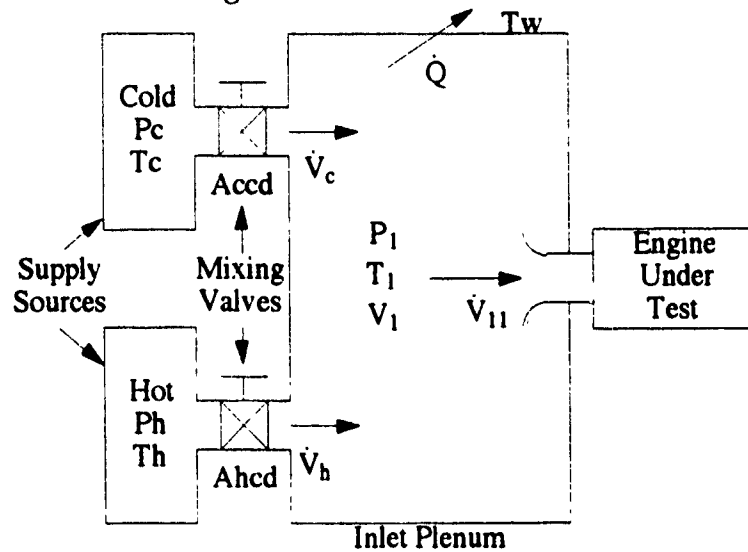


Figure 2: Schematic of Test Cell

The important parameters of the test cell are:

| | |
|-------------------|---|
| $P_1(t)$ | pressure in the inlet plenum |
| $T_1(t)$ | temperature in the inlet plenum |
| $\dot{V}_{11}(t)$ | volume flow rate into the test article |
| $\dot{Q}(t)$ | heat flow rate out of the inlet plenum |
| $T_w(t)$ | temperature of the inlet plenum wall |
| $\dot{V}_c(t)$ | volume flow rate of cold air |
| $\dot{V}_h(t)$ | volume flow rate of hot air |
| T_c | temperature of cold air supply (assumed constant) |
| T_h | temperature of hot air supply (assumed constant) |
| A_{ccd} | valve area ratio for cold air |
| A_{hcd} | valve area ratio for hot air |
| ΔP | difference between total and static bellmouth pressures |
| P_c | pressure of cold air supply (assumed constant) |
| P_h | pressure of hot air supply (assumed constant) |

Table 1 - Parameters

The nonlinear differential equations governing the system are:

$$P1(t) = \frac{\gamma P1(t)}{V1} [\dot{V}_h(t) + \dot{V}_c(t) - \dot{V}_{11}(t)] + \frac{\dot{Q}(t)}{V1} (\gamma - 1)$$

$$T1(t) = \frac{T1(t)}{V1} \left[\dot{V}_h(t) \left(\gamma - \frac{T1(t)}{T_h} \right) + \dot{V}_c(t) \left(\gamma - \frac{T1(t)}{T_c} \right) - \dot{V}_{11}(t) (\gamma - 1) + \frac{\dot{Q}(t)}{P1(t)} (\gamma - 1) \right]$$

$$\dot{Q}(t) = Ch \left(\frac{P1(t) \dot{V}_{11}(t)}{T1(t)^{1.76}} \right)^b As (Tw(t) - T1(t))$$

$$\dot{V}_{11}(t) = \frac{R T1(t)}{P1(t) - \Delta P} \frac{P1(t) A_{cd}}{\sqrt{T1(t)}} \sqrt{\frac{2 \gamma}{R (\gamma - 1)}} \sqrt{\left(\frac{P1(t) - \Delta P}{P1(t)} \right)^{\frac{2}{\gamma}} - \left(\frac{P1(t) - \Delta P}{P1(t)} \right)^{\frac{\gamma + 1}{\gamma}}}$$

$$Tw(t) = \frac{\dot{Q}(t)}{3600 C_m M}$$

$$\dot{V}_h(t) = \frac{R T_h}{P1(t)} \frac{P_h A_{hcd}}{\sqrt{T_h}} \sqrt{\frac{2 \gamma}{R (\gamma - 1)}} \sqrt{\left(\frac{P1(t)}{P_h} \right)^{\frac{2}{\gamma}} - \left(\frac{P1(t)}{P_h} \right)^{\frac{\gamma + 1}{\gamma}}}$$

$$\dot{V}_c(t) = \frac{R T_c}{P1(t)} \frac{P_c A_{ccd}}{\sqrt{T_c}} \sqrt{\frac{2 \gamma}{R (\gamma - 1)}} \sqrt{\left(\frac{P1(t)}{P_c} \right)^{\frac{2}{\gamma}} - \left(\frac{P1(t)}{P_c} \right)^{\frac{\gamma + 1}{\gamma}}}$$

where the meanings of the variables are given in the following table:

| | |
|----------|-------------------------------------|
| γ | ratio of specific heats |
| Ch | collection of constants |
| A_{cd} | bell-mouth throat area |
| b | empirical constant |
| As | inlet plenum effective surface area |
| C_m | heat capacitance of the metal |
| R | gas constant |
| M | mass of the metal |

Table 2 - Constants

Procedure:

Two different formulations of these differential equations were tried in order to strike a good balance of accuracy, utility, and computational complexity. Both formulations required extensive symbolic differentiation in order to derive and linearize the system model before implementing the Kalman filter. This symbolic differentiation was too extensive to handle by hand, therefore a symbolic mathematics package, Mathematica, was employed to compute the derivatives. Each formulation, however, followed the same basic procedure. [2] That basic procedure can be described as follows:

1. From the important parameters of the system model equations, select a set of these parameters (and their time derivatives) to be the state variables of the system. Select the remaining important parameters to be inputs. This results in the nonlinear state equations

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t))$$

where $\mathbf{x}(t)$ is the vector of state variables and $\mathbf{u}(t)$ is the vector of input variables. Also the vector of observations became

$$\mathbf{z}(t) = \mathbf{H}^T \mathbf{x}(t)$$

where the transpose of \mathbf{H} selects the observable state variables.

2. The equations for the time derivatives of the state variables define nonlinear differential equations describing the dynamics of the system. Discretize these equations using the Euler approximation to the derivative

$$\dot{\mathbf{x}}(t) \approx \frac{1}{T_s} [\mathbf{x}(t+T_s) - \mathbf{x}(t)]$$

This results in a set of nonlinear difference equations of the form

$$\mathbf{x}((k+1)T_s) = \mathbf{x}(k T_s) + T_s \mathbf{f}(\mathbf{x}(k T_s), \mathbf{u}(k T_s))$$

3. These nonlinear difference equations are then linearized about the state variables and the input variables to obtain a set of linear time-varying difference equations. This will be described in more detail later in this report.
4. The linearized difference equations are used to produce the extended Kalman filter.

The difference in the formulations derives from choosing different sets of state variables in step 1. above.

Results:

For the first realization, the state vector \underline{x} was selected as

$$\underline{x}(t) = \begin{bmatrix} P1(t) \\ T1(t) \\ Q(t) \\ V_{11}(t) \\ Tw(t) \\ \dot{V}_h(t) \\ \dot{V}_c(t) \end{bmatrix}$$

and the input vector \underline{u} was selected as

$$\underline{u}(t) = \begin{bmatrix} P_h(t) \\ P_c(t) \\ T_h(t) \\ T_c(t) \\ A_{hcd}(t) \\ A_{ccd}(t) \\ \Delta P(t) \end{bmatrix}$$

This formulation required that time derivatives of some of the original nonlinear equations be taken. In particular we took time derivatives of the equations for $\dot{Q}(t)$, $V_{11}(t)$, $Tw(t)$, $\dot{V}_h(t)$ and $\dot{V}_c(t)$ to obtain equations for $\ddot{Q}(t)$, $\ddot{V}_{11}(t)$, $\ddot{Tw}(t)$, $\ddot{V}_h(t)$, and $\ddot{V}_c(t)$. These derivatives resulted in very large and complex

nonlinear differential equations. The resultant vector of nonlinear differential equations, f , is given by

$$f(x(t), u(t)) = \begin{bmatrix} \dot{P}_1(t) \\ \dot{T}_1(t) \\ \ddot{Q}(t) \\ \dot{V}_{11}(t) \\ \dot{T}\dot{w}(t) \\ \dot{V}_h(t) \\ \dot{V}_c(t) \end{bmatrix}$$

These nonlinear equations had to be discretized and then linearized to form a set of linear state equations in order to apply the Kalman filter equations. [2] The discretization followed the approach in step 2 above. After dropping the T_s notation in the arguments of $\underline{x}(k T_s)$ and $\underline{u}(k T_s)$ we obtain

$$\underline{x}(k+1) = \underline{x}(k) + T_s f(\underline{x}(k), \underline{u}(k))$$

The linearization was carried out by taking derivatives with respect to \underline{x} to form a state transition matrix F and with respect to \underline{u} to form the input coupling matrix B . The linearized state equations were thus given by

$$\underline{x}(k+1) = F(k) \underline{x}(k) + B(k) \underline{u}(k)$$

where

$$F(k) = \frac{\partial}{\partial \underline{x}(k)} [\underline{x}(k) + T_s f(\underline{x}(k), \underline{u}(k))]$$

$$B(k) = \frac{\partial}{\partial \underline{u}(k)} [\underline{x}(k) + T_s f(\underline{x}(k), \underline{u}(k))]$$

The measurement equation was given by

$$z(k) = H^T \underline{x}(k) + v(k)$$

where

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and $\underline{v}(k)$ represents uncertainties (i.e. noise and/or error) in the measurements.

The Kalman filter variables, summarized in the table below, were given initial values so that the filter could run.

| | |
|-------------------|------------------------------------|
| S_p | predictive noise covariance matrix |
| S_c | corrective noise covariance matrix |
| H | output coupling matrix |
| R_n | output noise covariance matrix |
| Q_{rs} | input noise covariance matrix |
| \underline{z} | readings from actual test cell |
| \underline{x}_p | predictive state vector |
| \underline{x}_c | corrective state vector |

Table 3 - Kalman Filter Variables

The Kalman filter operated according to the following equations. The Kalman filter first calculated the Kalman gain L [2]

$$L(k) = S_p(k) H [H^T S_p(k) H + R_n(k)]^{-1}$$

The corrected noise covariance matrix S_c was then calculated according to [2]

$$S_c(k) = S_p(k) - L(k) H^T S_p(k)$$

The vector of measurements \underline{z} and the input vector \underline{u} were formed by reading the following values from the data files (of course, in a real-time environment these would come from the sensors directly).

$$\underline{z} = \begin{bmatrix} PIsim(k) \\ TIsim(k) \\ Twsim(k) \end{bmatrix}$$

$$\underline{u} = \begin{bmatrix} P_h \\ P_c \\ T_h \\ T_c \\ A_h(k) \\ A_c(k) \\ \Delta P(k) \end{bmatrix}$$

The pressures and temperatures of the supply sources P_h , P_c , T_h , and T_c are assumed to be constant.

The corrected state estimate $\underline{x}_c(k)$ (corrected using the measurements) was given by

$$\underline{x}_c(k) = \underline{x}_p(k) + L(k) [z(k) - H^T \underline{x}_p(k)]$$

The values from \underline{x}_c were used to evaluate F and B at time k since F and B are functions of $\underline{x}(k)$ and $\underline{u}(k)$. F and B were then used to calculate the one-step predicted noise covariance matrix $S_p(k)$.

$$S_p(k) = F(k) S_c(k) F^T(k) + B(k) Q_{ns} B^T(k)$$

The prediction of the next value of each of the state variables was made from the previously derived discrete nonlinear equations.

$$\underline{x}_p(k+1) = \begin{bmatrix} P1m(k+1) \\ T1m(k+1) \\ Qm(k+1) \\ \dot{V}_{11}m(k+1) \\ Twm(k+1) \\ \dot{V}_hm(k+1) \\ \dot{V}_cm(k+1) \end{bmatrix} = \underline{x}_c(k) + T_s f(\underline{x}_c(k), \underline{u}(k))$$

This process repeated over all the points in the simulation data.

The model from above was used with the derivatives of the equations for five of the state variables. The filter worked very well when simulated in the Mathematica package, but the equations for the F and B matrices were completely intractable for coding into FORTRAN (the equations describing the elements of the F and B matrices required 160 pages to print out), therefore a simpler model was developed. Instead of using \dot{Q} , \dot{V}_{11} , $\dot{T}w$, \dot{V}_h and \dot{V}_c as the state variables, their integrals Q , V_{11} , Tw , V_h , and V_c were used. This simplified the equations greatly, since the need for additional differentiation (to obtain $\ddot{Q}(t)$, $\ddot{V}_{11}(t)$, $\ddot{T}w(t)$, $\ddot{V}_h(t)$, and $\ddot{V}_c(t)$ in the previous formulation) was eliminated. The model worked well in this simplified

form, but had problems tracking the wall temperature T_w . This was apparently due to the fact that the diagonal element of S_p that allows correction from the sample of T_w was becoming very tiny, causing very little of the correction to affect T_w . When the engine began a maneuver, T_w would track poorly because the Kalman gain for T_w was too tiny to produce much correction. We solved the problem by adding a value of 0.5 to this element of S_p at every iteration. This resulted in a higher Kalman gain for T_w thus allowing the sample of T_w to provide more correction to the value of the predicted T_w in $\hat{x}_p(k)$. See the graphs provided in the appendix.

Conclusion:

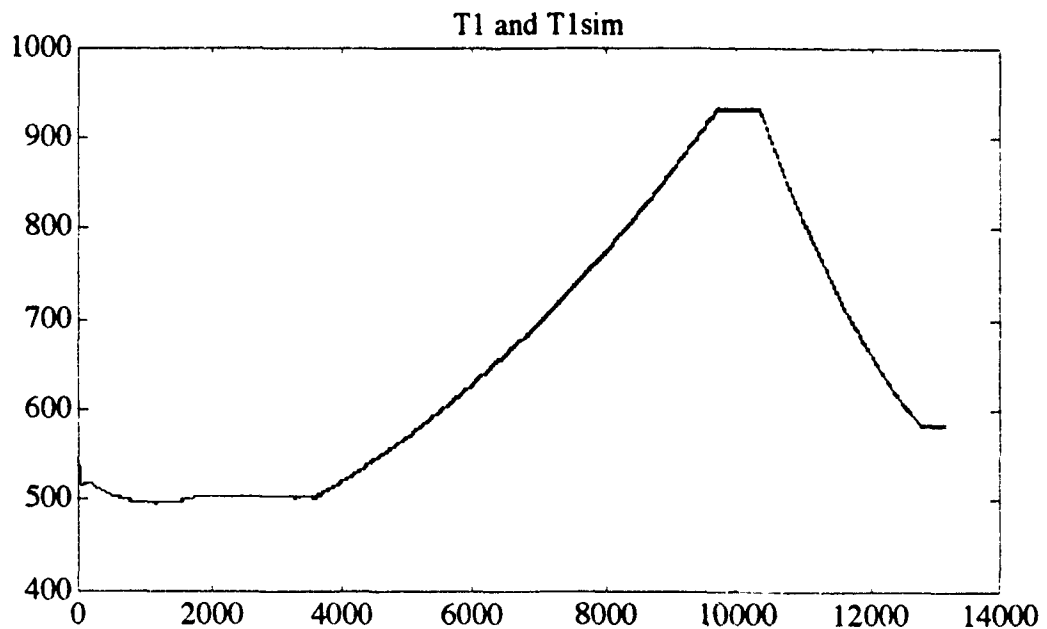
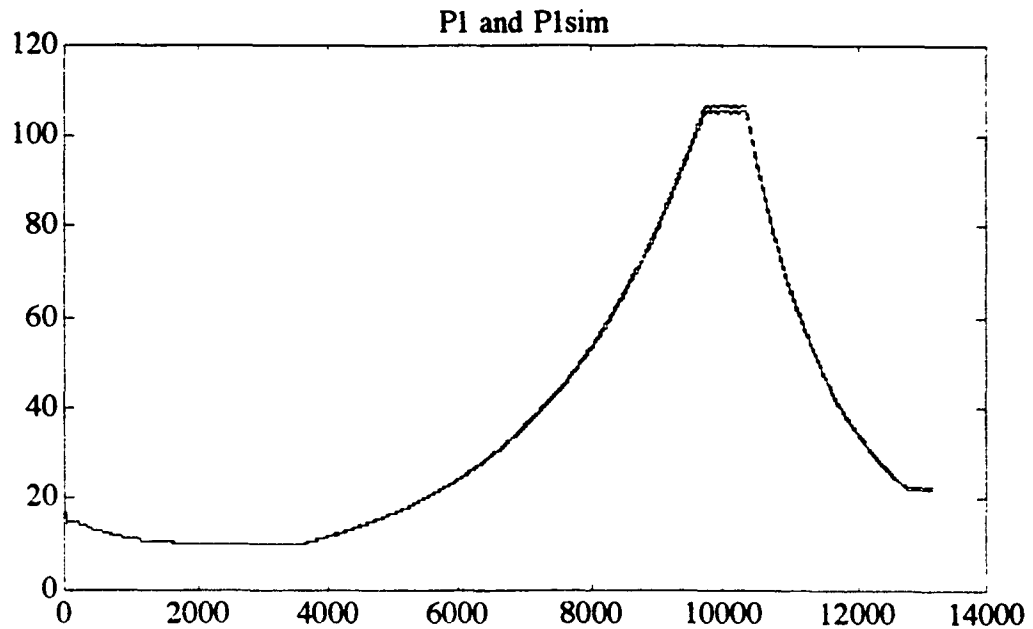
The developed extended Kalman filter tracked very well, but several topics need to be addressed. Future research should include the following topics:

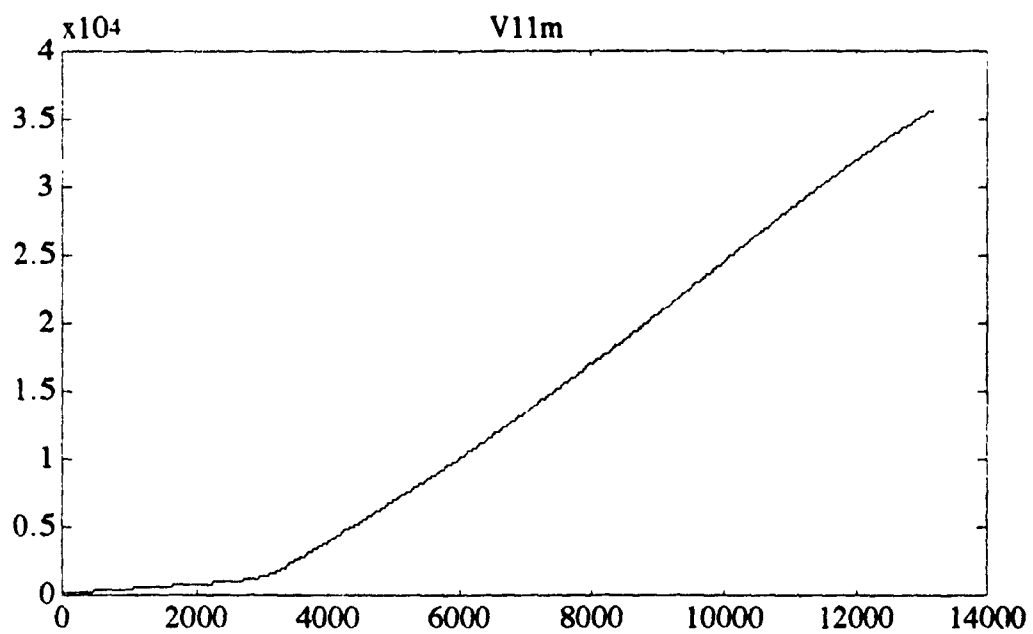
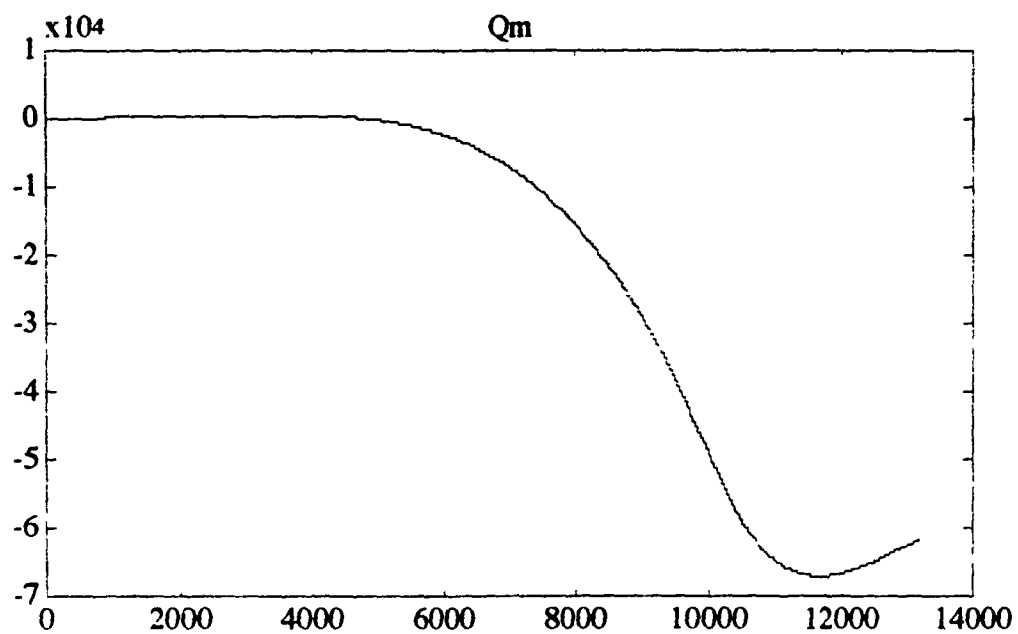
1. Integration of the FORTRAN implementation of the extended Kalman filter into existing simulation and control software to test performance more rigorously and evaluate whether the accuracy level is acceptable.
2. Evaluate whether the present implementation is fast enough for a real-time control situation.
3. Investigate faster implementations using multiprocessors and/or high speed processors.
4. Investigate the stability and sensitivity of the extended Kalman filter to disturbances and modelling error.

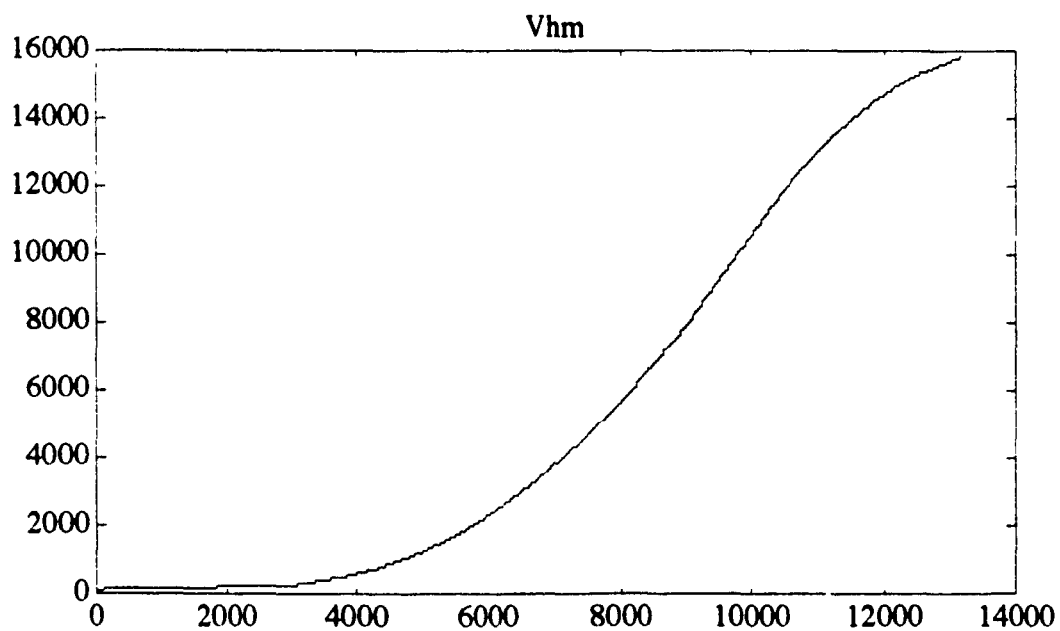
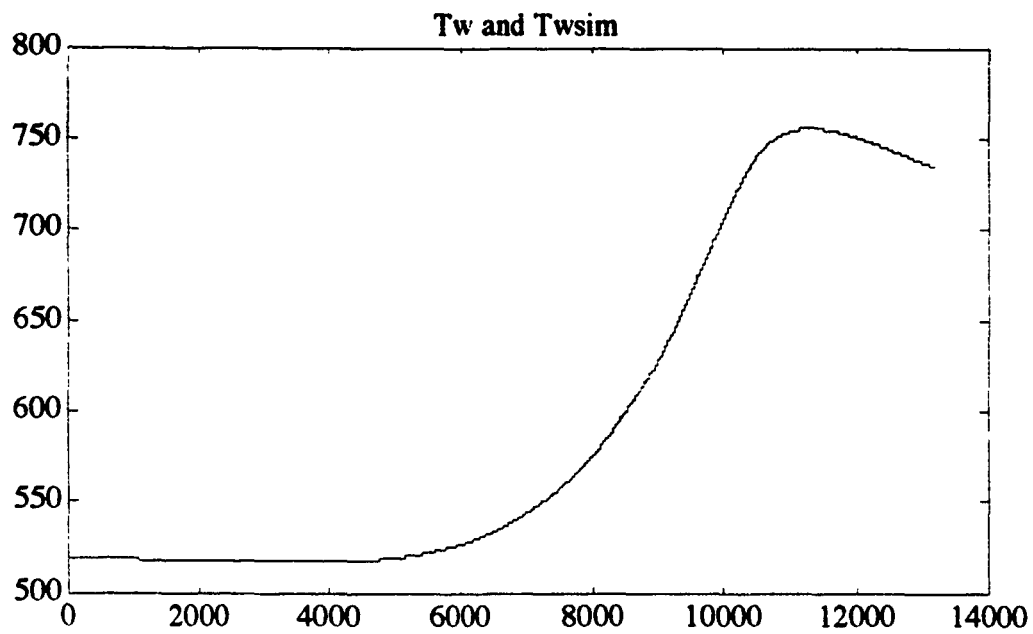
The results of this effort were very encouraging and strongly suggest that the use of an extended Kalman filter can improve the accuracy of the inverse-process model in the model-following controller.

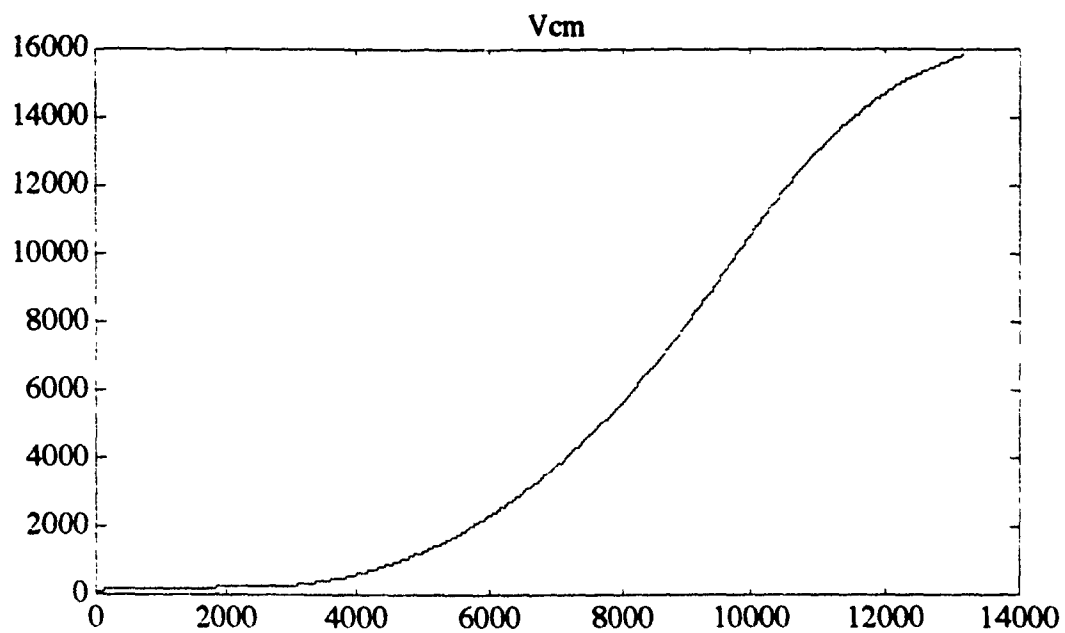
Appendix

The graphs below show the actual data from P1, T1, and Tw plotted with the predicted values of P1, T1, and Tw. The plots for Q, V₁₁, V_h, and V_c show the behavior expected of the integral of the quantities of interest.









References

- [1] M. J. Chaney and J. J. Beaman, "Comparison of Nonlinear Tracking Controllers for a Compressible Flow Process", submitted to *ASME Journal of Dynamic Systems, Measurement and Control*
- [2] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control*, Prentice-Hall, Edgewood Cliffs, N.J., 1971

**PARALLEL SIGNAL PROCESSING
FOR
TURBINE ENGINE TESTING**

BY

Theodore A. Bapty
Ben A. Abbott

SEE: BEN A. ABBOTT

DEVELOPMENT OF A COMPUTATIONAL MODEL OF THE
LASER-INDUCED FLUORESCENCE (LIF) OF HYDROGEN (H₂) WITH
EXCITATION AT 193 NANOMETERS

Clinton Daniel Benefield
Graduate Student
U.T. Space Institute, Tullahoma, TN

Abstract

The development and application of LIF diagnostics for O₂ and NO using excimer lasers is in progress at AEDC. The potential exists for LIF of H₂ using 193 nm. excitation. A theoretical/computational model is needed to accurately predict locations and structure of absorption and emission transitions.

Introduction

Several experiments reported since 1978 indicate the potential for LIF of hydrogen using excitation at 193 nanometers. Development and application of LIF diagnostics of oxygen and nitric oxide is in progress at the Arnold Engineering Development Center, but AEDC also needs H₂ diagnostics for many applications.

A good computational model is needed for the following: 1. to predict the locations and structure of absorption and emission transitions; 2. to provide a capability to study fluorescence signal dependence on H_2 density and temperature, collisional quenching, and photoionization; 3. to provide a comparison with experimental measurements; 4. to provide an estimate of signal level and feasibility of application without the need for extensive laboratory measurements; 5. to aid in analysis of flow-field measurements.

Considerable research of the hydrogen molecule and its isotopes has been done in the past seventy years. I have extensively researched the literature, and I will summarize the material most relevant to my research. Also, I will discuss the theoretical and experimental work that has been done on this subject at AEDC this past summer.

Background

For the proposed LIF-diagnostics of H_2 , the excitation involves a two-photon absorption from the ground state, labeled $X\ ^1\Sigma_g^+$, to the excited state labeled (E,F) $\ ^1\Sigma_g^+$. This state then fluoresces, resulting in a transition to the $B\ ^1\Sigma_u^+$ state [see Figures 1 and 4]. The Σ represents the state with $\Lambda = 0$, where Δ

is the orbital angular momentum component along the internuclear axis. The superscript "1" represents a singlet ("3" would be a triplet). These numbers refer to the multiplicity. The "g" and "u" represent "gerade" or "ungerade" (even or odd state). These symbols are important for determining which transitions are possible, as described in Hund's cases (c.f. G. Herzberg, Spectra of Diatomic Molecules). The notation used in labeling the E and F states can be confusing in the literature because different authors label levels differently. The E and F states are two different states, but their potential energy curves are at nearly the same energy and have an avoided crossing [see Figure 2]. Many authors used a combined EF state, so that when they label a vibrational number, it will be labeled v_{EF} or simply v , where the levels of the E and F states (or of the inner minimum and the outer minimum) are counted sequentially according to their energy. Other authors label them separately as v_E and v_F , or as v_1 and v_2 . For example, $v_{EF} = 6$ is $v_E = 2$.

In 1934, O.W. Richardson published his research in a book titled Molecular Hydrogen and Its Spectrum.¹ This book contributed significantly to the study of the hydrogen molecule, and most papers on H_2 from 1935 through the 1960's referred to this book. In his book, Richardson discusses some quantum mechanics, and he analyzes his data in order to properly label the

vibrational and rotational levels. Some of his labels later were found to be incorrect, especially in regard to the E,F state.

Much improvement on Richardson's data and labeling was made by Gerhard Dieke, who reported his work on hydrogen and its isotopes until his death in 1965. Dieke's experimental data are still the most complete tables in existence, so I will discuss some of his most significant work as it applies to my thesis. In 1936, he reported his study of the transition from $2s\ ^1\Sigma$ to $2p\ ^1\Sigma$ [see Figure 3], which he also labeled 2A and 2B.² These are the states we now refer to as the E and B states, using Herzberg's notation. These transitions lead to infrared radiation. Dieke observed some irregularities for higher vibrational levels, and proposed another state at approximately 3500 cm^{-1} above his 2A state. He called the perturbing state $(2p\sigma)^2$, also called 2K by Dieke, and labeled the F state by Herzberg. Dieke later found transitions attributed to this state, which he reported in a paper in 1949.³

Dieke reported what may be his most significant paper in 1958. In this paper, Dieke included extensive tables of the term values he had found experimentally. Only relative intensities on a scale from 0 to 10 were given, but he did label the term values for 50 states of hydrogen and its isotopes.⁴ Dieke's data were later compiled by H.M. Crosswhite and printed in a book in 1972

titled The Hydrogen Molecule Wavelength Tables of Gerhard Heinrich Dieke.

In 1960, McLean, et. al., provided a complete listing of the theoretical calculations that had been reported on the hydrogen molecule from 1927 to 1960.⁵ The first paper they report is W. Heitler and F. London in 1927.⁶ Another paper of significance is H. James and A. Coolidge in 1933.⁷ Their paper is significant because their method was used in 1960 by W. Kolos and C.C.J. Roothaan, which began a decade of increasingly more accurate and complicated calculations which are most significant to my work. Kolos and Roothaan used a 40-term James and Coolidge type function in their calculations of the electronic wave functions of H_2 .⁸ This function converges increasingly more slowly for increasing internuclear separation R , so that the most accurate calculations are for $R \leq 2.0$ a.u. Later that year, E.R. Davidson reported that his calculations indicated that the potential curves of the E and F states form a double-minimum because of an avoided crossing.⁹

In 1961, Davidson published his own calculations of the $^1\Sigma_g^+$ state using a 20-configuration wave function.¹⁰ He identified an inner minimum at $R=1.9$ a.u. as $1s\sigma 2s\sigma$ $^1\Sigma_g^+$, and an outer minimum at $R=4.3$ a.u. as $(2p\sigma)^2$ $^1\Sigma_g^+$. Davidson's calculation employed a wave function without explicit dependence on the interelectronic distance, but his results were

not accurate enough to make a detailed comparison between theory and experiment.

Robert Boyle reported a calculation in 1968 which used a 24-term variational wavefunction expressed in elliptic coordinates.¹¹ He compares his calculation to the 1961 Davidson paper, and also to a paper by Davidson and Rothenberg (1966), which had more accurate results for the inner minimum using 45-50 configurations. Boyle's results represent a significant improvement over Davidson's results. However, Kolos and Wolniewicz state that "they still fail to account for a sizable portion of the correlation energy, especially in the vicinity of the outer minimum."¹²

In the 1969 paper quoted above, Kolos and Wolniewicz report the following:

All previous results were obtained in the Born-Oppenheimer approximation, i.e., by completely neglecting the coupling between electronic and nuclear motion. For the state under consideration one may expect significant deviations from the Born-Oppenheimer approximation, especially in the region of the avoided crossing of the potential energy curves.¹²

In their calculation, a 50-term wavefunction was used at $R=1.4$ a.u., a 63-term wavefunction at $R=4.3$ a.u., and a 51-term wavefunction at $R=8.0$. Their results are only slightly better

than Davidson and Rothenberg's for the inner minimum, but became "progressively significantly better" for the outer side of the maximum.

Also of significance is their analysis of the calculated transitions and the experimentally observed ones. They state the following:

The most interesting qualitative feature of the results ... is certainly the prediction of additional vibrational levels located in the outer minimum below those whose existence could be inferred from the experimental data. Thus, for H_2 , the lowest transition directly observed from the outer minimum was that corresponding to $v = 7$ ($v_2 = 4$) level in our notation. ...no experimental evidence has been found for the existence of still lower vibrational levels.

As will be seen, these first four vibrational levels of the F state were not experimentally observed until the 1980's.

In 1977, Wolniewicz and Dressler reported their study of the EF and GK $^1\Sigma_g^+$ states of hydrogen, using an adiabatic calculation of vibronic states.¹³ They state that "the vibrational levels of the $^1\Sigma_g^+$ states show nonadiabatic behavior due to coupling between the electronic and vibrational motions." Wolniewicz reported in 1976 his calculated nonadiabatic

corrections to the rotational energies and rotational constants B_v computed using a variational perturbation method for several vibrational levels of H_2 , HD, and D_2 .¹⁴ Again in 1979, Wolniewicz, et.al., reported calculations of the nonadiabatic coupling of the EF and GK states.¹⁵

From 1985-1987 significant improvements on both experimental and theoretical studies of H_2 were made by P. Senn et.al. In 1985, they identified the lowest rotation - vibration levels of the $(2p\sigma)^2 F^1\Sigma_g^+$ state.¹⁶ In particular, they identified the $v_F = 0-4$ and $J = 0-5$ levels, which had been predicted by Wolniewicz but had not been found experimentally by Dieke. Their measurements are more accurate than Dieke's and correspond very well with the theoretical predictions of Wolniewicz and Dressler, who made improved calculations in 1985.¹⁷

In this paper, they report computations of the EF, GK, and $H\bar{H}$ of hydrogen using a 129, 118 and 110 term wavefunction, respectively. This calculation provided a more continuous approximation for the double minimum, whereas their previous computations had used different wavefunction expansions for different R values.

In 1987, Senn and Dressler reported observations of tunneling resonances in the rotation-vibration structures of the

EF double-minimum state of H_2 and D_2 .¹⁸ They compare their work with the previous theoretical and experimental term values. They report in another paper in 1987 on a new study of the EF state, this time with spectroscopic identification of rovibronic levels above the potential barrier ($v_{EF} = 9 - 20$).¹⁹ These extend the spectral data known for these states by several vibrational quantum numbers. Also, they identify perturbations for high v values (e.g. $v=19$) from the $I^1\Pi_g^+$ state and GK $1\Sigma_g^+$ state.

Studies using laser-induced fluorescence were first published in the late 1970's. In 1978, Kligler and Rhodes reported selective excitation of the E,F $1\Sigma_g^+$ state of H_2 by two-photon absorption at 193 nm, with fluorescence arising from $1\Sigma_g^+ - B^1\Sigma_u^+$ transitions.²⁰ Their report includes some information regarding collisional quenching and radiative decay. Their experimentally determined values for the cross-section of two-photon excitation compares favorably with theory. Also, they state that "other collisional processes, such as rotational and vibrational relaxation, electronic energy transfer, para-ortho conversion, and formation of HD by four-center reaction may also be studied by this technique."

In 1980, Kligler, et. al. reported on their more thorough study of the collisional and radiative properties of the E,F state

of hydrogen.²¹ They used an ArF laser in these experiments, with an untuned linewidth of about 170 cm^{-1} centered at 193.4 nm . They could tune the laser from 192.6 to 194.2 nm . They observed near-infrared emission, which they attributed to a transition from the $v=2$ level of the inner minimum to $v=0$ and $v=1$ of the $B\ ^1\Sigma_u^+$ state. Emission originating from $J'=0, 1$, and 2 is observed, but they say $J'=3$ is not excited. They state that "while the electronic relaxation processes we have examined are quite rapid, at higher laser intensities the fastest loss process--and hence the limiting factor in achievable excitation density--is photoionization of the excited molecules. While the characteristic signature of a two-photon process is a signal which grows as the square of the laser intensity, when photoionization is prevalent, the signal 'saturates' and grows only linearly with intensity; meanwhile, the laser energy is increasingly channeled into ion production."

In 1983, Huo and Jaffe report on their calculation of the two-photon absorption cross section of the transition from the X state to the E,F state.²² They refer to the experiments of Kligler et.al. and disagree with the labeling of the vibrational levels. They state, "In the experiments of Kligler et. al., the final state is at the $v=6$ level. This was labeled as $v=2$ in Sharp's compilation, based on the progression of a vibrational series localized in the inner well. Sharp's notation was used by Kligler

et.al. After considering the characteristics of the vibrational wavefunction and its transition moments, we feel that the $v=2$ label is a misnomer." Their report includes some useful diagrams and tables.

Some of the more recent papers on LIF of H_2 report on: ionization-dissociation, photoelectron angular distributions, vibrational branching ratios, absorption cross sections, unusual fluorescence, alignment of EF state, and two-dimensional imaging of molecular hydrogen in diffusion flames. Some of this work was done for a slightly different wavelength region than that which we are (or have been) most interested in. Also, much of the experimental setup differed from the one used at AEDC. However, some of these papers do have graphs indicating some rather sharp peaks in their spectral data, indicating that LIF can be done with reasonable effectiveness for the transitions we are interested in at AEDC.

Pummer et.al. report on vacuum ultraviolet stimulated emission of H_2 , with two-photon excitation using a 193-nm ArF laser.²³ They report that "at 300 K the relative equilibrium rotational state populations in the ground $X^1\Sigma_g^+$ state for $J=0,1,2$ and 3 levels are 1.0, 4.9, 0.83 and 0.52, respectively." The $B^1\Sigma_u^+$ state has a radiative lifetime of about 600 ps (the exact values depend on vibrational quantum number), with a

collisional quenching rate $k_B = 1.3 \times 10^{-9} \text{ cm}^3/\text{s}$ at 300 K. The E,F state has a radiative lifetime of 90 ns with a collisional quenching rate of $k_E = 2.1 \times 10^{-9} \text{ cm}^3/\text{s}$. The C $^1\Pi_u$ state and inner minimum of the E,F state nearly coincide.

Their experimental work involved 193 nm ArF laser with effective power of 2 GW with pulse duration of 10 ps. The laser radiation had a bandwidth of 5 cm^{-1} and divergence of $10 \mu\text{rad}$. They used pressures from 20 - 500 Torr, observed transitions from E to B at 922.2 nm, 837.0 nm, 754.41 nm, and observed a transition at 1121.0 nm for pressure >600 Torr.

Results

This past summer, I began my investigation of hydrogen and LIF. I will summarize the progress I have made.

First, I studied the papers given to me regarding LIF, and began a library research for the papers referred to in those most recent papers. Also, I studied G.Herzberg's Spectra and Structure of Diatomic Molecules. My main purpose at that time was to become familiar with the notation used for molecules, such as labels for different electronic states, labels for different branches (e.g. P, Q, R), and Hund's rules.

Having studied Herzberg and after finding several papers

on H_2 , I then began my own calculations of the term values for the E,F, and ground (X) states of hydrogen. I used the constants listed in Huber and Herzberg's Constants of Diatomic Molecules.

The formula I used is given as follows:

$$T = T_0 + [\omega_0(v + 1/2) - \omega_0 x_0(v + 1/2)^2 + \omega_0 y_0(v + 1/2)^3 + \dots] + \{ [B_0 - \alpha_0(v + 1/2)]J(J+1) - [D_0 + \beta_0(v + 1/2)]J^2(J+1)^2 + \dots \}.$$

I compared my calculated values with the experimental values of Dieke. My calculations for the ground state were very nearly the same as Dieke's data. However, for the E and F states, my values were considerably different and increasingly worse for higher vibrational and rotational quantum number.

After doing more research, I learned that the E,F state is irregular due to perturbations between the two states. Also, because so little was known about the four lowest vibrational levels of the F state before the 1980's, the values for B_v were good only for $v_F=4$. It became apparent then that a different approach was needed in order to calculate the transitions. It was then that I thoroughly investigated the theoretical work of Kolos, Wolniewicz, and other theoreticians already mentioned previously.

Using Dieke's experimental data, I calculated the wavelengths we could expect to find. We did some experimental work to see if any of my predictions were correct. We graphed

the excitation spectrum on a computer, and attempted to determine the wavelengths [see Fig. 5]. I did find several likely lines that matched up with the approximate locations of our lines, although there is a discrepancy on the left side of our graph. We show two lines in the neighborhood of 192.85 nm, but my calculations indicated only one line at that point. The other lines seem to match up extremely well with expected transitions from $v''=0$ (X state) to lower vibrational and rotational levels in the E and F states. However, the more recent experimental data of Senn et.al. with the newly discovered $v_F = 0-4$ may lead to transitions near 193 nm that I did not calculate from Dieke's data.

An emission spectra graph is also given [see Fig. 6]. However, there is yet considerable question as to which transitions we are seeing. These graphs were made from data taken on an older spectrometer which had severe limitations in its calibration and in the excess noise which distorted the real data. Better spectrometers were not available at the time. I do know the approximate wavelengths of the largest peaks. These include peaks at approximately 755 nm, 775 nm, and 865 nm.

Conclusion

Considerable theoretical work has been done on the hydrogen molecule, especially with regard to potential energy curves, adiabatic calculations with nonadiabatic corrections, absorption cross sections, collisional and radiative properties, and ionization-dissociation. Experimental work using LIF is becoming increasingly important for the study of H_2 , with many applications in the study of combustion, extraterrestrial atmospheres, vacuum-uv sources, and laser-based analytical probes.

With all the progress that has been made in the past decade on the study of hydrogen, a considerable body of information has become available for application at AEDC. However, there are specific applications for which more information is required, but which have not been investigated thoroughly or sufficiently. In particular, we have need for a computational model which can accurately calculate the transitions and the relative intensities, if not the actual intensities, for various temperatures and pressures.

The work that has been done so far on H_2 , particularly with regard to the E,F state, has used a variety of pressures at room temperature. The experimental conditions which will be used at AEDC differ from most, if not all, of the previous work on

H₂. Miles et.al. have applied LIF in their work at Princeton with temperatures at 800 K, but the application of LIF of H₂ with excitation at 193 nm is still very recent, with considerable room for improvement in the accuracy of prediction and in the precision of experimental measurements.

With considerable work remaining for me to develop a reliable computational model of the LIF of H₂ with excitation at 193 nm, I have submitted this project as a thesis proposal and am in the process of writing a mini-grant proposal.

Diagrams

The figures appear as follows:

1. Potential energy curves of significant states discussed in this paper.
2. Potential curve for the avoided crossing of E and F states.
3. Diagram of electronic states of the H₂ molecule taken from

Herzberg's Spectra of Diatomic Molecules;

4. Energy level diagram of two-photon excitation.
5. Experimental excitation spectrum.
6. Experimental emission spectra.

BIBLIOGRAPHY

1. O.W. Richardson, Molecular Hydrogen and Its Spectrum, Yale University Press, New Haven (1934)
2. G.H. Dieke, Phys. Rev. **50**, 797 (1936)
3. -----, Phys. Rev. **76**, 50 (1949)
4. -----, J. Mol. Spect. **2**, 494 (1958)
 - H.M. Crosswhite, ed. The Hydrogen Molecule Wavelength Tables of Gerhard Heinrich Dieke. Wiley Interscience (1972)
5. A.D. McLeon, A. Weiss, M. Yoshimine, Rev. Mod. Phys. **32**, 211 (1960)
6. W. Heitler, F. London, Z. Physik **44**, 455 (1927)
7. H. James, A. Coolidge, J. Chem. Phys. **1**, 825 (1933)
8. W. Kolos, C.C.J. Roothaan, Rev. Mod. Phys. **32**, 219 (1960)
9. E.R. Davidson, J. Chem. Phys. **33**, 1577 (1960)
10. -----, J. Chem. Phys. **35**, 1189 (1961)
11. R.J. Boye, J. Mol. Spect. **26**, 36 (1968)
12. W. Kolos, L. Wolniewicz, J. Chem. Phys. **50**, 3228 (1969)
13. L. Wolniewicz, K. Dressler, J. Mol. Spect. **67**, 416 (1977)
14. L. Wolniewicz, J. Mol. Spect. **63**, 537 (1976)
15. K. Dressler, R. Gallusser, P. Quadrelli, L. Wolniewicz, J. Mol. Spect. **75**, 205 (1979)
16. P. Senn, P. Quadrelli, K. Dressler, G. Herzberg, J. Chem. Phys. **83**, 962 (1985)
17. L. Wolniewicz, K. Dressler, J. Chem. Phys. **82**, 3292 (1985)
18. P. Senn, K. Dressler, J. Chem. Phys. **87**, 1205 (1987)
19. P. Senn, K. Dressler, J. Chem. Phys. **87**, 6908 (1987)
20. D.J. Kligler, C.K. Rhodes, Phys. Rev. Letters **40**, 309 (1978)
21. D.J. Kligler, J. Bokor, C.K. Rhodes, Phys. Rev. A **21**, 607 (1980)
22. W.M. Huo, R.L. Jaffe, Chem. Phys. Lett. **101**, 463 (1983)
23. H. Pummer, H. Egger, T.S. Luk, T. Srinivasan, C.K. Rhodes, Phys. Rev. A **28**, 795 (1983)

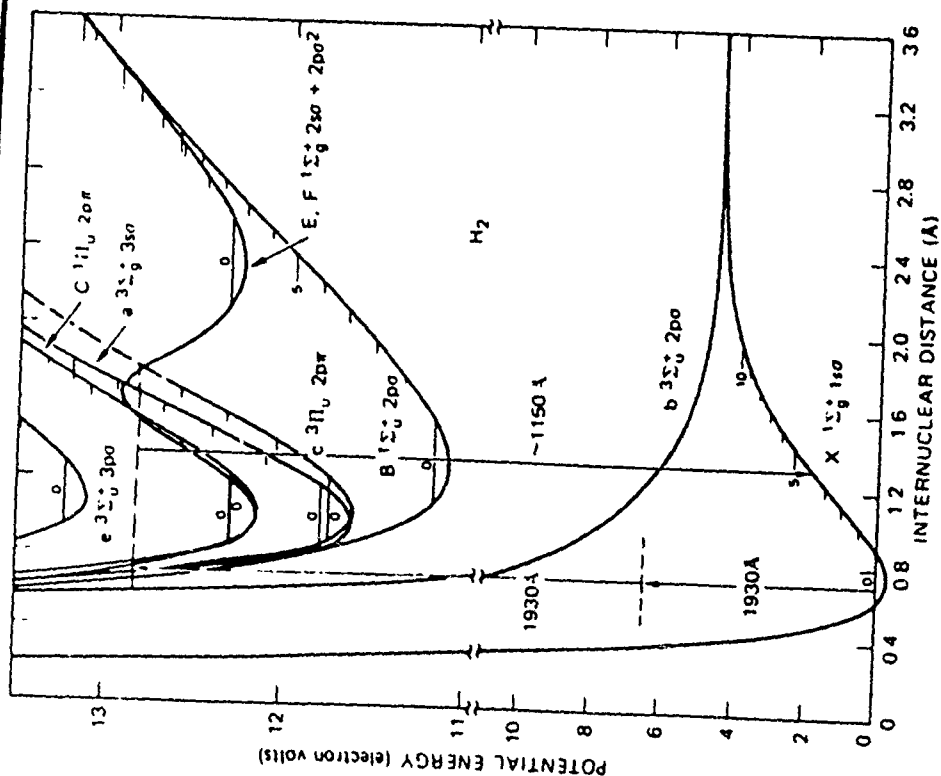


FIG. 1. Energy-level diagram of H_2 showing the X, B, C, and E, F states of interest in this experiment, as well as other states in this energy range, based on potential curves in T. E. Sharp, At. Data 2, 119 (1971). Note the change in vertical scale above 10 eV.

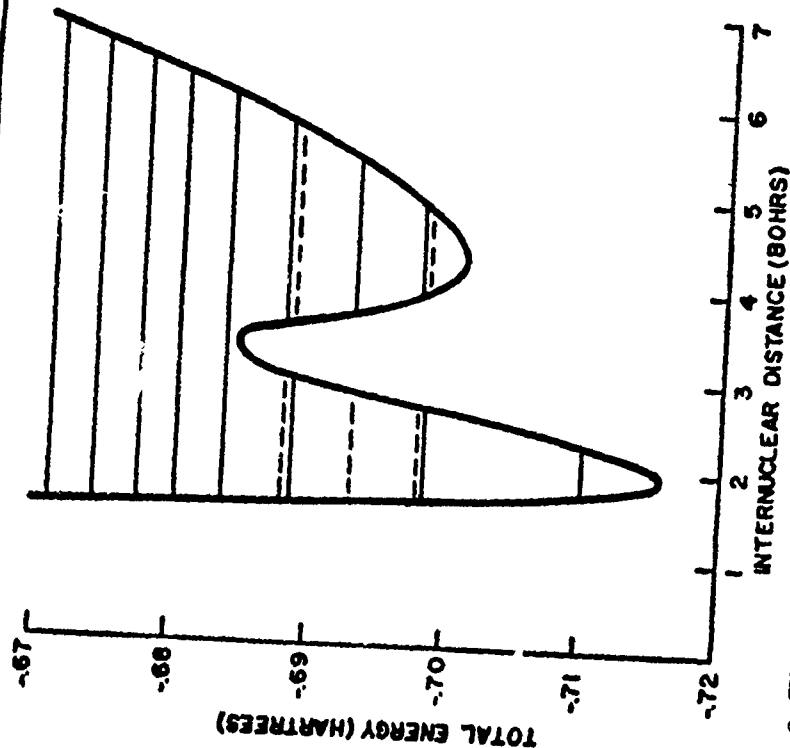


FIG. 2. The potential curve and vibrational levels for the first excited Σ^+ state of the hydrogen molecule. The term levels are indicated by solid lines in regions in which the probability density is large.

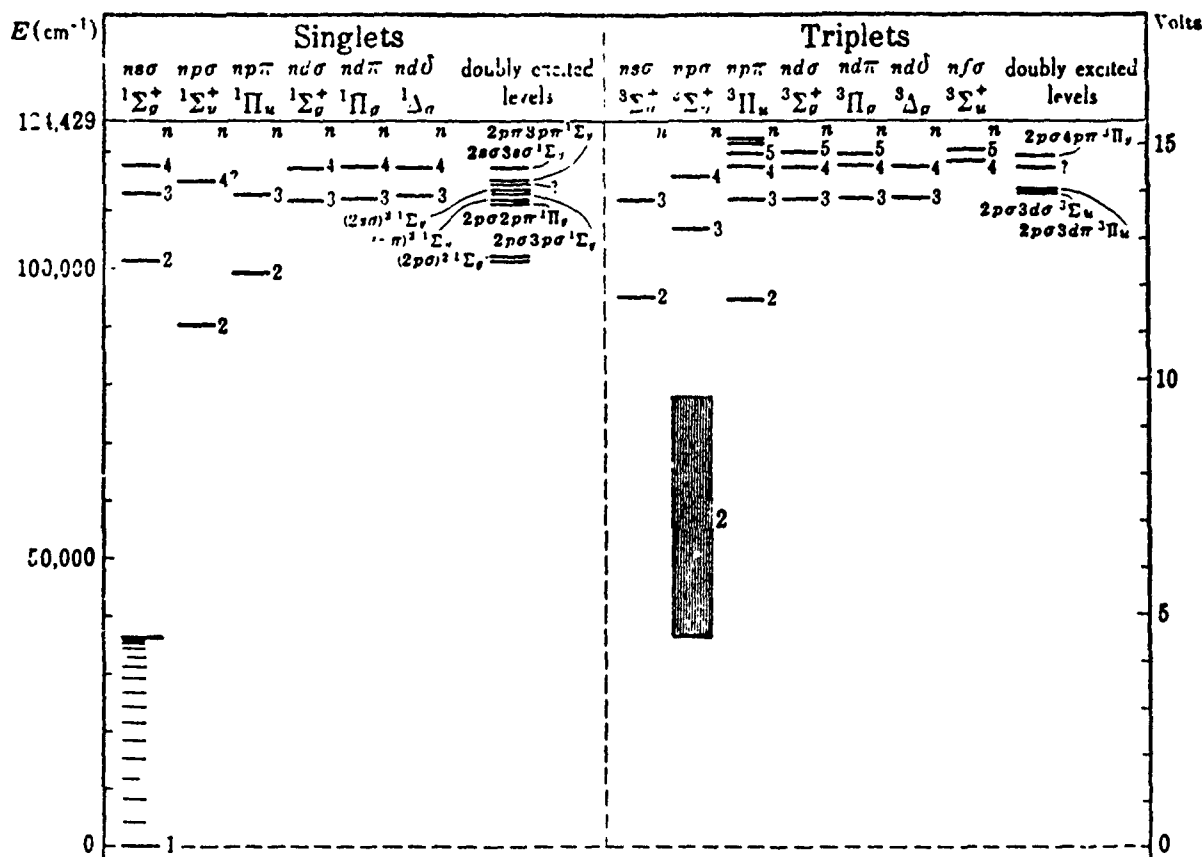


FIG. 3 . Diagram of the Observed Electronic States of the H_2 Molecule. The data for most of the ordinary states are taken from Richardson (32), and those for the doubly excited states from Richardson (586), and Richardson and Rymer (587). For the ground state, the observed vibrational levels are also indicated (shorter horizontal lines). Above each column (with the exception of the doubly excited states) the orbital of the excited electron and the term type is given. The numbers beside the levels are the n values. The state $2p\sigma^3\Sigma_u^+$ is the unstable state resulting from two normal atoms. The corresponding continuum is indicated by hatching. It should, of course, extend up to infinity.

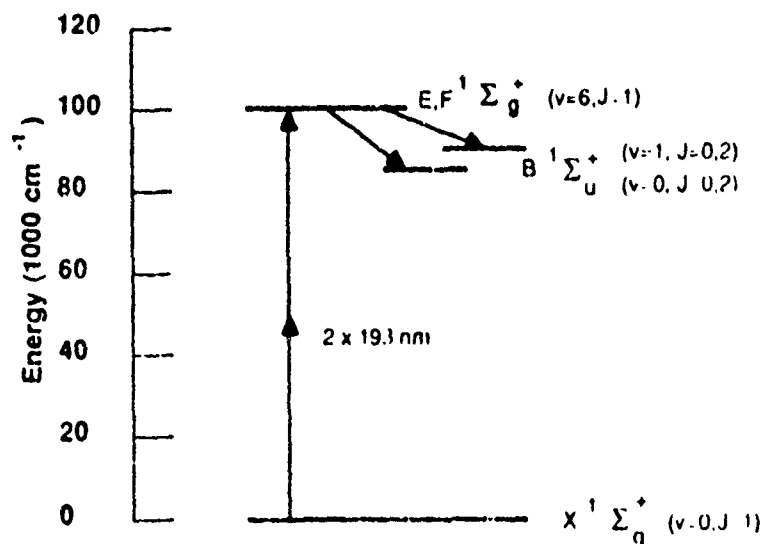


Fig. 4. H_2 energy-level diagrams for the relevant states for two-photon I.F.

Hydrogen Excitation Spectrum, H2072401

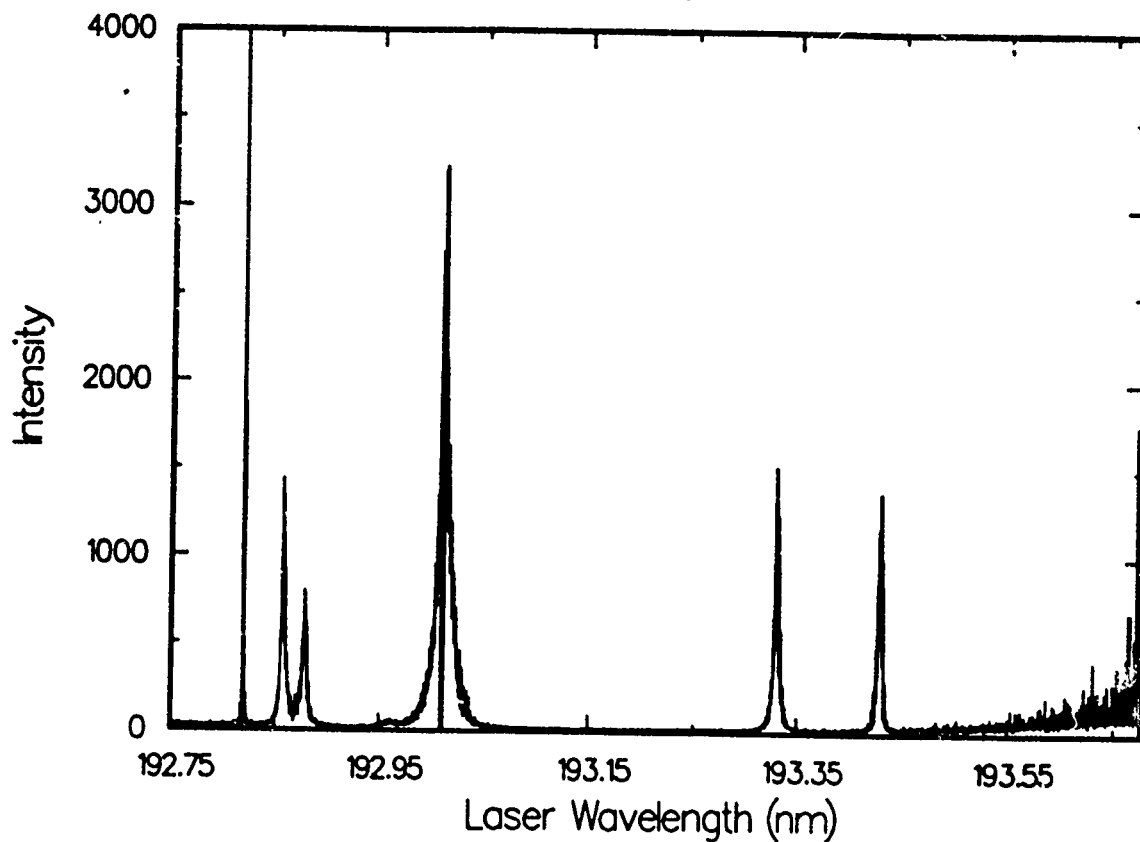


Fig. 5

H₂ Emission Spectrum, 100 torr, GG-335 Filter

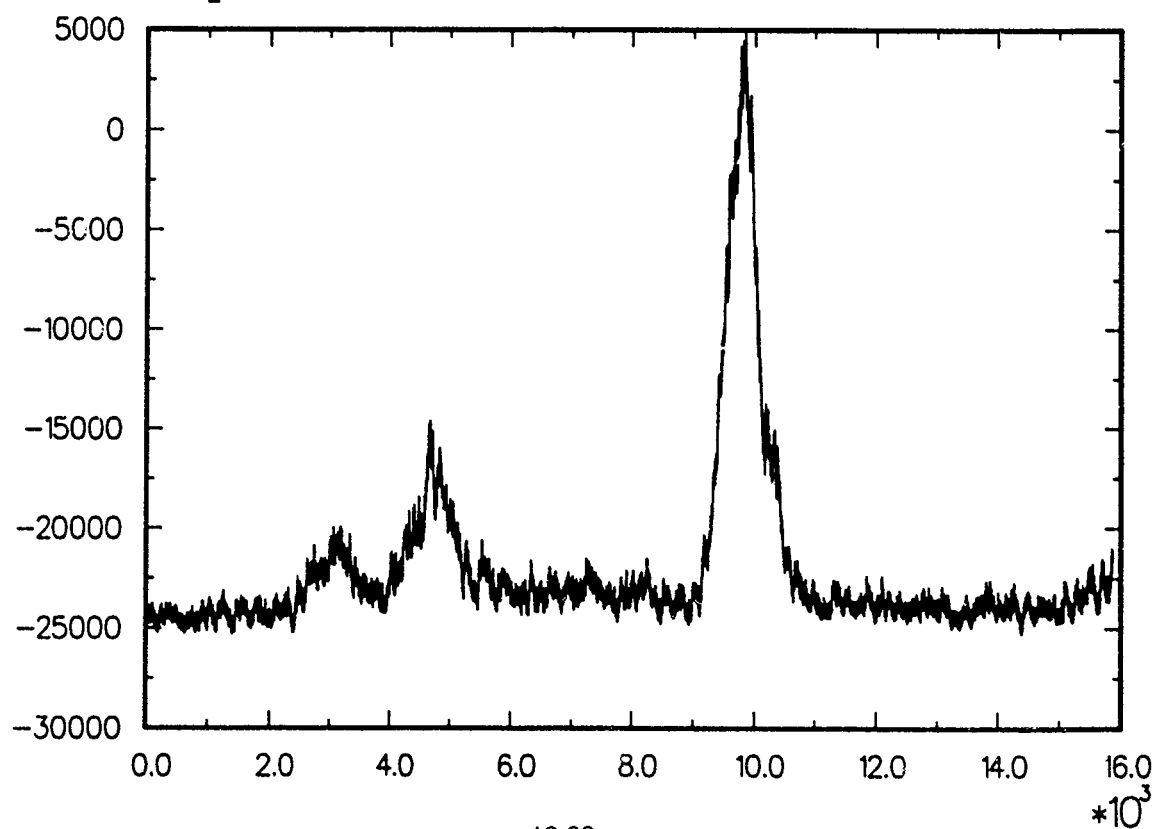


Fig. 6

PLIF DEVELOPMENT FOR WAKE PHYSICS ANALYSIS

AFOSR Summer Research Program

Final Report

August 16, 1991

John V. Dempsey, Jr.

Arnold Engineering & Development Center (AEDC)
Propulsion Wind Tunnel Building
Arnold AFB, Tennessee 37389

ABSTRACT

The following report describes the author's assignment at Arnold Engineering Development Center, Arnold Air Force Base, Tullahoma, Tennessee under the AFOSR Summer Research Program. The author participated in an active Air Force sponsored technology project titled, "PLIF Feasibility for Wake Physics". The project is being performed by Calspan Corporation/AEDC Division, operating contractor for aerospace flight dynamics testing at AEDC. The project is being conducted to investigate the feasibility of exploiting laser induced fluorescence to indicate the presence of selected gas molecules in the wake of a hypervelocity projectile in the hypervelocity gun range at AEDC. A Nd:YAG Pulsed Dye Laser and a CCD array camera produce an image of the downstream wake of the hypersonic upper atmospheric vehicle to provide data for development of mathematical models.

Related concepts include instantaneous measurements of number density, velocity, hypersonic flow, pulsed laser, array camera, and shock tube.

The accounts will be presented in an introductory manner. The start of the internship, coincided with the planning stage for the project. Ten weeks later, the intern term, the findings are being presented here.

INTRODUCTION

Planar Laser Induced Fluorescence (PLIF) is a technique whereby a known gas, naturally occurring or injected into the flow, is fluoresced by a short duration high energy pulse of laser radiation. Different molecules in the flow excite under different frequencies of the laser. The light is formed into a sheet by cylindrical optical elements. While the molecules fluoresce in the sheet, a cross-sectional view of a vehicle's wake is captured by an array camera. PLIF is a noninvasive technique. It will not disturb the flow field characteristics.

PURPOSE

The accounts represented are the findings for the Air Force technology project, " PLIF Feasibility for Wake Physics". The purpose of the project is to develop a qualitative measurement of the molecular number density within the wake of a hypersonic vehicle in a planar cross-section normal to the trajectory.

Vehicles at hypersonic speeds develop wakes that are underexpanded. An underexpanded wake represents the

condition where the base pressure is more than the atmospheric pressure. As the wake pressure decreases to match atmospheric conditions, the wake flow expands. As the wake expands, the supersonic velocity increases, resulting in rapid changes in flow parameter values with respect to time.

The project calls for the capture of a planar picture of the wake in the range of 30 to 100 model radii downstream of the model base. This planar picture will quantitatively describe the number density distribution of the selected gas molecule in the wake expansion. Due to a minimal funding allowance, the project must be incorporated with an existing project scheduled for Range G, the AEDC hypervelocity gun range. The type of vehicle tested is immaterial in relation to the purpose of the project.

EXPERIMENTAL PROCEDURES

Evolution of the project incorporates six stages. The stages include planning, build-up, laboratory check-out, installation, testing, and final analysis.

The initial stage, planning, combines available expertise, materials, and resources, with the design goal. The goal is to present the Air Force Technology Division with a qualitative electronic photograph, in the form of an

array, of a selected plane in the vehicle's wake. At AEDC, a Spectra Physics, DCR-4 Nd:YAG Pumped Dye Laser system is used for PLIF research and development. The Nd:YAG Laser is tunable for wavelengths 220-1064 nanometers. Different gas molecules become fluoresced as the frequency of the laser beam is varied. The intensity of the molecular fluoresce quantitatively indicates the number density. An electronic camera records the fluorescent radiation from the molecules and produces a two dimensional array of the digital intensity values.

The build-up stage enables the engineer to assemble the equipment for the test. This step includes the procurement of materials. Due to a limited materials budget, on-hand hardware was used for the design when possible. Figure 1 depicts the overview set-up in Range G. As shown, the Pulsed Spectra Physics Nd:YAG Laser is tuned from its initial frequency of 1.06 microns to approximately 226 nanometers. While in the system, green crystals are injected to the developing laser beam. Although any color may be used, for this system and design, visibility is increased with the use of a green dye. The green laser beam is collimated by the sheet forming optics and passed through the clear SuprasilTM window (Fig. 2). A high quality window must be used instead of regular glass due to the radiation wavelength and power of the non-visible beam.

The thin sheet is projected across the tunnel such that it intersects with the projectile's intended path. An aluminized front surface mirror with a magnesium fluoride overcoat is mounted just off the projected trajectory in an orientation such that the CCD camera can view the model wake through another port (Figs. 1 and 3). When tuned to approximately 226 nanometers, the laser sheet fluoresces the nitric oxide (NO) molecules, a gas generated by the physical processes in the high velocity wakes, and it is captured by a quick exposure of the CCD camera. A narrow band pass filter admits only 250 nanometer light, the wavelength of the NO fluorescence, into the CCD camera. The data, fluoresced molecules of NO gas, will be stored by a computer, and then analyzed upon completion of the test runs.

The laboratory timing check-out is a critical step in the operation of the project. A typical cone shaped model, with 3 cm base radius, can pass the station in Range G at a rate of 16,000 fps (.488 cm/microsec). As previously stated, the project calls for the capture of the wake at a downstream distance of 30 to 100 model radii. At this speed, the time elapsed from the model passage to the shot for capture of the wake must be accurately timed to an order of 10 microseconds. Precision timing tests have been performed in the laboratory, using a flat flame burner to simulate the wake. The Range G engineers routinely include

time sequences for the various schlieren stations of the gun range. The real success of the project will require incorporating Range G expertise with the optics technology, PLIF.

The installation and test stages are strongly dependent on Range G scheduling. Since the test must be incorporated with an existing test scheduled in Range G, the PLIF experiment will not be exercised until October 1991 (Fiscal Year 92). The installation will include the mounting of the suspended mass for supporting the turning mirror for the CCD array camera (Fig. 5). The mount (Fig. 4) will be welded to the inside wall of the Range G tank. It will support the mass by three non-flexible steel cables and will be attached as illustrated. The angle of the mirror must be accurately set to correctly reflect the sheet image into the camera. The author designed both the mount and the suspended mass for stress, safety, and movement for the extreme dynamic conditions of the gun range. The laser system, lenses, and the array camera will be installed outside the shock tunnel and mounted from the floor.

The final analysis will hopefully present the Air Force with a qualitative array of the wake specie density values determined by PLIF. If the project goals are satisfied, the project will provide the Air Force with an

additional diagnostics tool for future research endeavors. PLIF can be an accurate measurement system undertaking new improvements constantly. The PLIF technique can significantly increase the knowledge concerning the development and operation of hypersonic upper atmospheric vehicles.

RESULTS

The results have not been obtained as of the printed date of the report. The experiment procedure has been completed through the laboratory timing stage. The integrated design has been established and the materials have been either previously obtained, fabricated, or requisitioned. All systems are ready for installation for the October 1991 test.

A potential source of error is developed by any variance of the suspended mass' (turning mirror) location and angle in relation to the model flight path prior to the activation of the electronic camera. Analysis shows that if the mirror moves axially by 1 cm, either upstream or downstream, the camera will lose approximately 1.6 cm on one side of the 5 cm x 5 cm square target. If the model propagates 1 cm from its intended flight path, the target captured by the array camera will be cut by 20 percent.

CONCLUSIONS and FORESIGHT

As stated in the purpose, the camera captures the data, fluoresced molecules of NO gas, of the expansion in the form of an array. The next question or step is to determine the exact location of the image. This step, although obtainable, is beyond the scope of the current project, and therefore, not required.

Quantitative information can be extracted, but when the photographed array captures the planar view of the wake, the picture will be skewed due to the turn angle of the line of sight, known as the Keystone effect, as shown in Figure 6. This is also beyond the scope of the project. An algorithm must be created to correct the wake photo before a modeling process can be designed.

Another future area of research could incorporate multiple planar imaging stations using PLIF, as shown by Figure 7, in a fashion similar to the shadow graph system of Range G.

ACKNOWLEDGMENTS

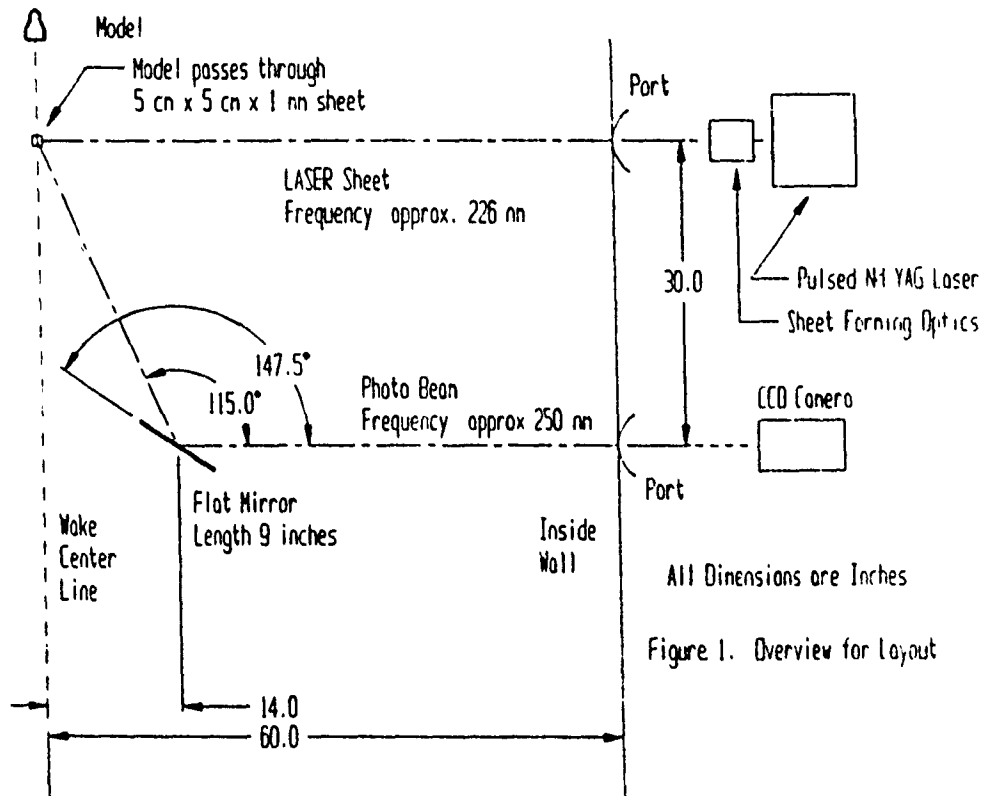
The author would like to thank the Air Force Office of Scientific Research, Research and Development Laboratories, of Culver City, California, and Major D. C. Hart, of Air Force Department of Technology, Arnold AFB, for their judicious investment. The author would also like to acknowledge Fred L. Heltsley, of the Calspan Corporation, AEDC Division, for his invaluable leadership and guidance as the internship mentor.

REFERENCES

Lozano, A., van Cruyningen, I. J., Hanson, R. K., 1990, "Planar Laser Induced Fluorescence Scalar Measurements in a Turbulent Jet", Presented at the Fifth International Symposium on Application of Laser Techniques to Fluid Mechanics, July 9 - 12, 1990, Lisbon, Portugal.

Hanson, R. K., 1989, "Advanced Diagnostics for Reacting Flows", AFOSR-TR-89-1433

ILLUSTRATIONS



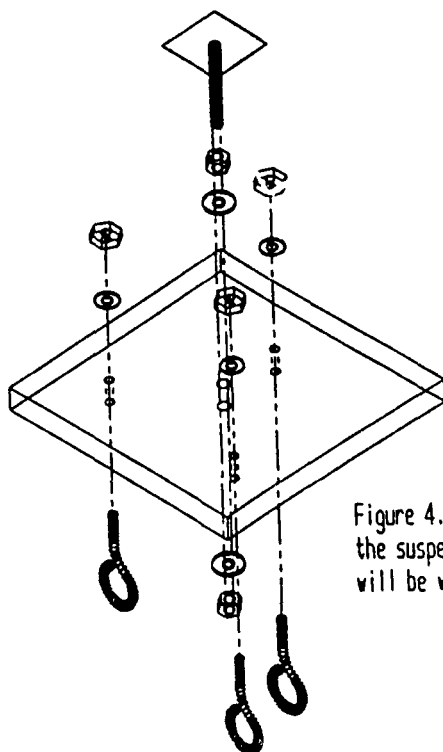


Figure 4. Design of the nount that supports the suspended mass in Range G. The stud will be welded to the inside wall.

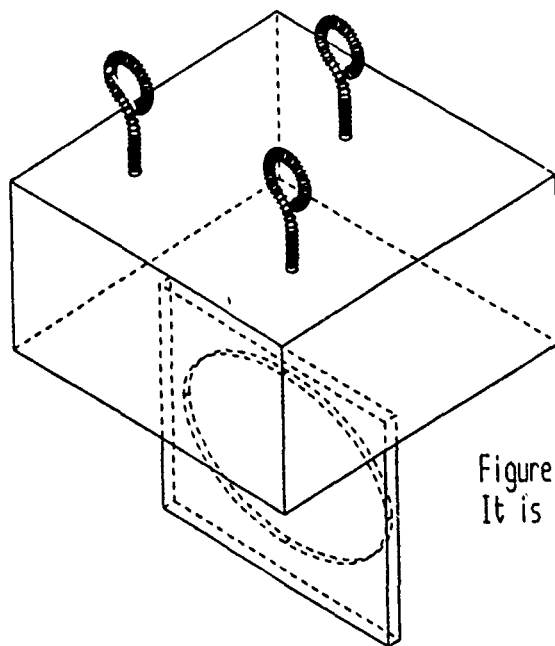
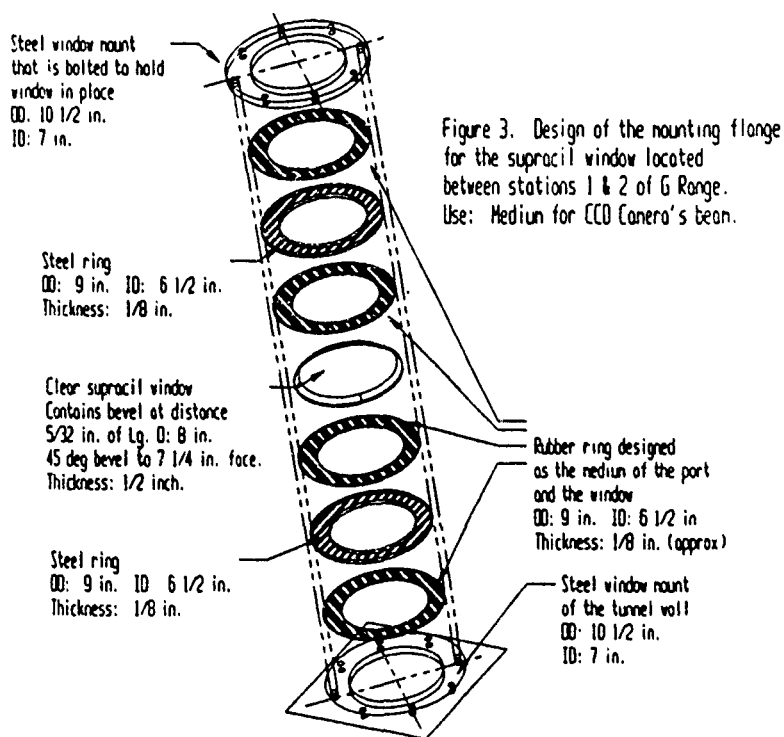
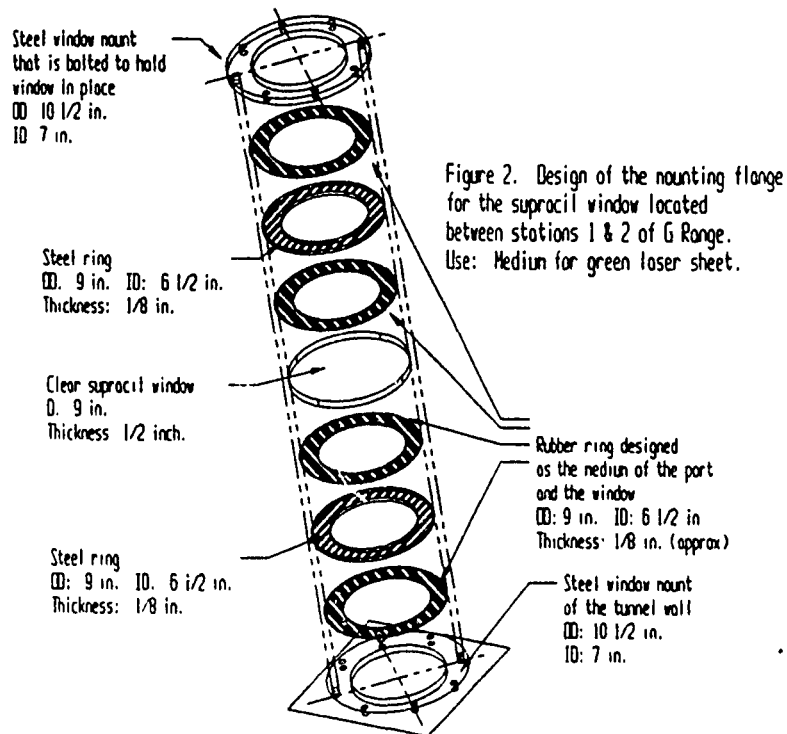


Figure 5. Mirror nount design. It is to be suspended in Range G.



Research & Development Laboratories

FINAL REPORT

PARAMETRIC EVALUATION OF FINITE ELEMENT
MODELING TECHNIQUES

| | |
|--------------------|---|
| Prepared by: | Dollena S. Hawkins |
| Academic Rank: | Master's Degree |
| Department and | Mathematics Department |
| University: | Tennessee Technological University |
| Research Location: | Turbine Engine Division, AEDC, Arnold Air Force Base, TN |
| USAF Researchers: | Major J. A. Davis Mr. J. S. Schroeder |
| Date: | August 15, 1991 |

PARAMETRIC EVALUATION OF FINITE ELEMENT
MODELING TECHNIQUES

by

Dollena S. Hawkins

ABSTRACT

Algor's finite element analysis software was used to model basic structural geometries. Parametric evaluation of finite element modeling techniques was achieved by generating models with different element and mesh densities and comparing the model generated data to theoretical or experimental results to determine model accuracy.

ACKNOWLEDGEMENTS

Thanks is extended to the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Also, Research & Development Laboratories must be mentioned for their assistance in all administrative and directional aspects of this program.

A special "thank-you" is extended to everyone in DOPT for making me feel like a part of the group. In addition, gratitude is also expressed to Major J. A. Davis, Mr. J. S. Schroeder, Mr. K. Nichols, Mr. J. Cromer, and Mr. K. Johnson for their invaluable knowledge and assistance in the completion of this project.

I. INTRODUCTION:

Finite element analysis (FEA) is often used in the design and study of structures. While FEA does not provide exact results, "engineers assume that models include enough elements in critical areas to accurately reflect loading, boundary conditions, material properties, and geometry." [Reference 5]

Accuracy of results, however, is a function of many variables including element formulation and numerical techniques used in the code, mesh size, and mesh construction. While many FEA programs provide automatic mesh generation and matrix bandwidth optimization, the skill and expertise of the person generating the model still factors into the accuracy of the model results. Accordingly, every FEA model has some degree of error. Optimization and error analysis are two areas of interest for both the designer and researcher.

The Mesh Density Handbook being developed at AEDC is a fundamental approach to error analysis. Parametric evaluation of finite element modeling techniques is achieved by generating models with different element and mesh densities and comparing the model generated data to theoretical or experimental results to determine model accuracy.

II. OBJECTIVES:

While the outline and format for the Handbook have been developed, minimum data has been collected. The objective of this research was to determine the effect of mesh density on model accuracy for five basic geometries:

- 1) Thin-walled cylindrical pressure vessel

- 2) Semi-infinite flat plate with a center hole
- 3) Semi-infinite flat plate with u-notches
- 4) Semi-infinite flat plate with v-notches
- 5) Semi-infinite flat plate with fillets
- 6) Curved beam with applied force on end

While AEDC has FEA packages including ANSYS, NASTRAN, and SAP-6, which run on a mainframe computer, a FEA package which runs on a 386 PC was selected for this project. There were two reasons for this. One, the PC software will be used by the Analysis Department which will be using the handbook. Two, access to the PC did not require a security clearance while access to the mainframe did.

III. LITERATURE REVIEW:

Algor Interactive Systems markets the software used in this research. References 2 and 4 provide an overview of the Algor FEA package which is summarized here. The software consists of modules which provide modeling and pre-processing, analysis, and post processing and visualization. Modeling and pre-processing is completed in VisiCad Plus, three modules of Algor's system. Models are created in the drawing module, SuperDraw II. After adding boundary conditions and constraints, the model is transferred to the Decoder module which creates the input file for the processor. Within the Decoder, material and element properties are defined for the model. In SuperView, the third module, hidden line and shaded pictures of the models can be viewed. After Decoding, the model's input file is submitted to the appropriate processor within Algor's analysis system. The SuperSap module offers a wide

variety of processors including those for static stress, weight, and center of gravity calculations. Other processor modules from Algor provide dynamic analysis of models. For this project, the static stress processor was utilized.

The Algor FEA processor calculates stresses in a model at nodal points, rather than at the centroid. Since stresses are calculated at nodal points, the Algor system offers a precision index display. When a node is shared by more than one element, a stress value for that node is created for each element. A precision value is given for that node by averaging the maximum and minimum values of the node from the associated elements. A precision index of zero is given if the node is attached to only one element. The benefit of this precision index is that it enables the designer to evaluate the accuracy of the analysis and refine the mesh density in areas of low precision.

After analysis is complete, the output is written to a file and is also available in visual form. SuperView provides both original and displaced views of the model.

IV. APPROACH & RESULTS

When the models were created, the following guidelines were used

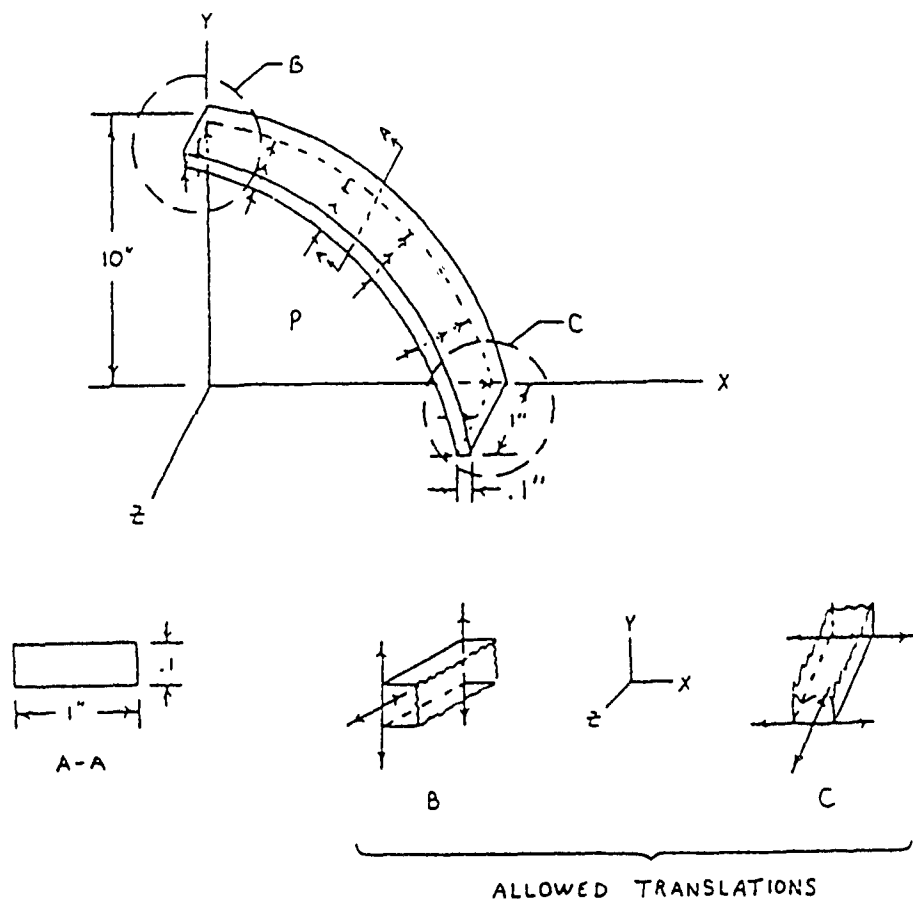
There are certain rules of thumb . . . (in FEA) . . .
(One guideline states that for most accurate results, elements should be as nearly equilateral as possible. When using rectangular elements the aspect ratio (longest side:shortest side) should not be greater than 3:1. Another rule states that angles between adjacent sides should not be less than 30 degrees, or greater than 120 degrees. Also mesh density should be highest in regions where high stresses are expected. [Reference 2]

A. Thin-Walled Cylindrical Pressure Vessel

Due to symmetry, only a quarter section of a thin-walled cylindrical pressure vessel was modeled. Interfacing edges were restrained as shown in Figure 1. The Algor program calculated deflection for applied internal pressure was compared to a theoretical deflection for varying mesh densities. Calculated hoop stress was also compared to a theoretical hoop stress. Using Algor's quadrilateral plate element, the quarter section was modeled with 2, 4, 8, 16, 32, and 64 elements. The theoretical deflection and the theoretical hoop stress were calculated using the strength of materials classical solution found in Reference 3. The deflection ratio (calculated/theoretical) and the hoop stress ratio (calculated/theoretical), each as a function of the number of elements of the quarter section of a cylindrical pressure vessel are shown in Figures 2 and 3.

B. Semi-Infinite Flat Plate with a Center Hole

Again, due to symmetry a quarter section of the geometry was modeled. Interfacing edges were restrained as shown in Figure 4. The program calculated stress for applied axial tension was compared to a theoretical elastic stress for varying mesh densities. The computer solution was obtained by using 2-D quadrilateral elements. The quarter section of the circular hole was divided into 4, 8, 10, 12, and 16 elements. Theoretical stress was calculated using the formula from Reference 6 for a central circular hole in a member of rectangular section. The stress ratio (calculated/theoretical) as a function of the division of the circle is shown in Figure 5.



Notes: 1) NOMINAL RADIUS = 10"
 OUTSIDE RADIUS = 10.05"
 INNER RADIUS = 9.95"

2) $P = 1000 \text{ LBS}$

3) $E = 30 \times 10^6 \text{ PSI}$
 $\nu = .3$
 $T = 70^\circ \text{ F}$

4) ALGOR DECODER PARAMETERS APPLIED:

ELEMENTS: TYPE: PLATE

ELEMENTS: INFO: FORMATION: REDUCED SHEAR

ELEMENTS: GROUP: STEEL MATERIAL PROPERTIES FROM ALGOR'S LIBRARY

ELEMENTS: COLOR: THICKNES = 1

TREF = 70

PRESSURE = 1000

PPx = 0

PPy = 0

PPz = .5

ANALYSIS: STATIC

GLOBAL: LOAD CASE: A=1, B=0, C=0, D=0

DECODE: ALL, INTERSECT LINES

DEFAULT VALUES USED FOR ALL OTHER PARAMETERS

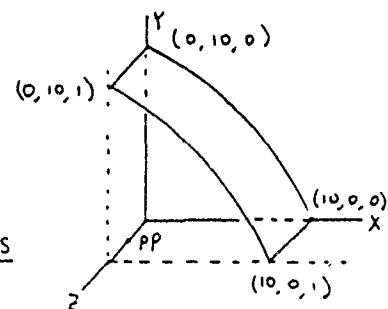


FIGURE 1

Cylindrical Pressure Vessel: Deflection

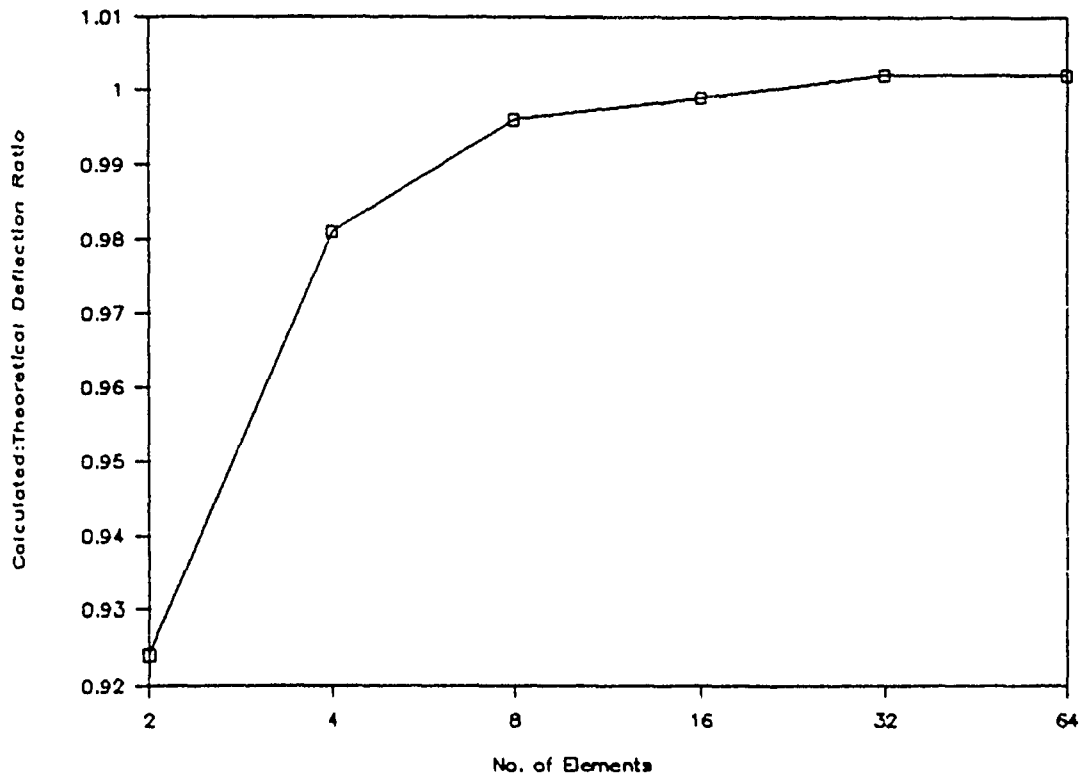


FIGURE 2

Cylindrical Pressure Vessel: Stress

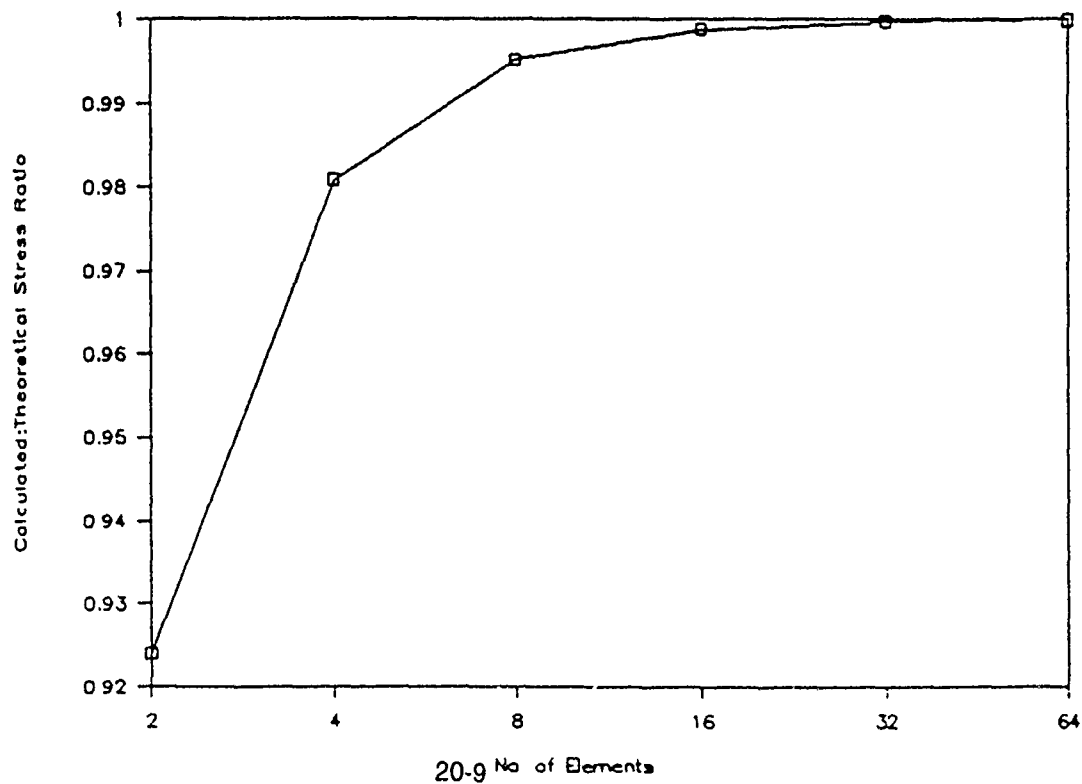
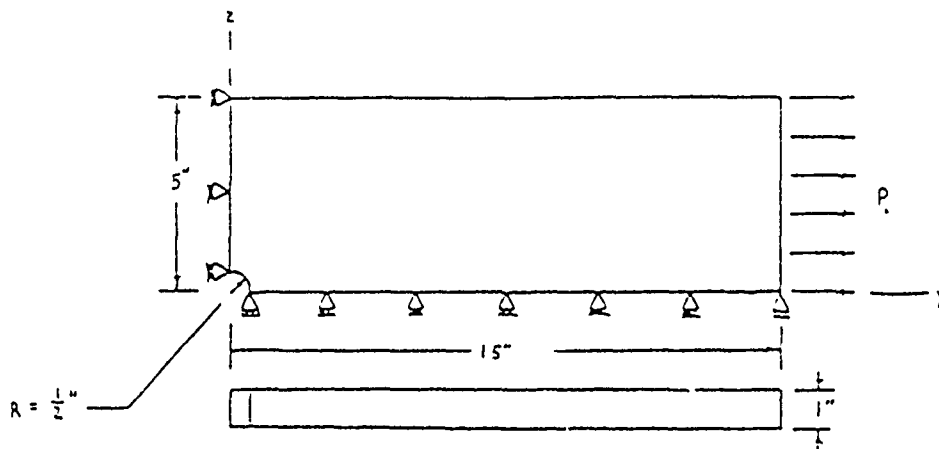


FIGURE 3



- Notes: 1) $P_i = \frac{5500}{N}$ WHERE N EQUALS THE NUMBER OF NODES ON THE INDICATED EDGE
- 2) BOUNDARY CONDITIONS APPLIED AT EACH NODE ALONG EDGES
- 3) $E = 30 \times 10^6$ PSI
 $\nu = .3$
 $T = 70^\circ$ F
- 4) ALGOR DECODER PARAMETERS APPLIED:
 ELEMENTS: TYPE: 2-DIM
 ELEMENTS: INFO: TYPE: PLANE STRESS
 ELEMENTS: GROUP: STEEL MATERIAL PROPERTIES FROM ALGOR'S LIBRARY
 ELEMENTS: COLOR: THICKNES = 1
 $TREF = 70$
 ANALYSIS: STATIC
 GLOBAL: LOAD CASE: $A=0, B=0, C=0, D=0$
 DECODE: ALL, INTERSECT LINES
 DEFAULT VALUES USED FOR ALL OTHER PARAMETERS

FIGURE 4

Plate w/Hole: Stress

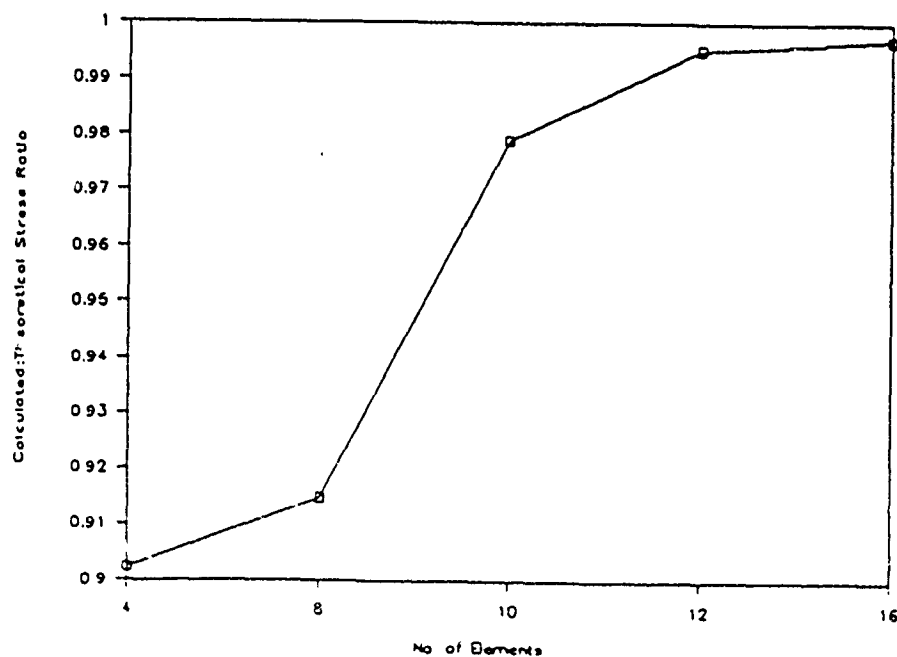


FIGURE 5

C. Semi-Infinite Flat Plate with U-Notches

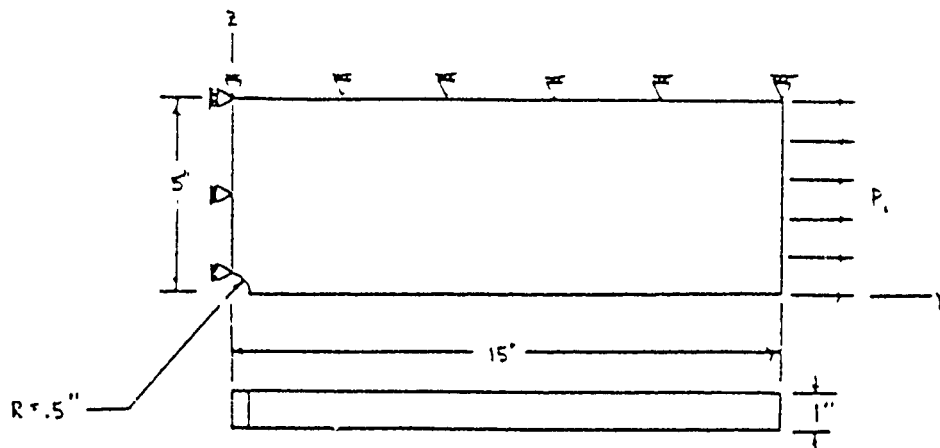
A quarter section of the geometry was modeled with interfacing edges restrained as shown in Figure 6. The maximum stress at the notch as a function of the number of elements used to model the curved edge of notch was evaluated and compared to a theoretical stress. The plate was modeled using 2-D quadrilateral elements. The notch was divided into 4, 8, 10, 12 and 16 sections. The theoretical stress was calculated using the formula for two u-notches in a member of rectangular section found in Reference 6. The stress ratio (calculated/theoretical) as a function of the division of the curved edge of the notch is shown in Figure 7.

D. Semi-Infinite Flat Plate with Two V-Notches

A quarter section of the geometry was modeled with interfacing edges restrained as shown in Figure 8. The maximum stress at the notch as a function of the number of elements used to model the curved edge of notch was evaluated and compared to theoretical stress. The plate was modeled using 2-D quadrilateral elements. The notch was divided into 3, 6, 9, and 15 sections. The theoretical stress was calculated using the formula for two v-notches in a member of rectangular section found in Reference 6. The stress ratio (calculated/theoretical) as a function of the division of the curved edge of the v-notch is shown in Figure 9.

E. Semi-Infinite Flat Plate with Fillets

A half section of the geometry was modeled with interfacing edges restrained as shown in Figure 10. The program calculated stress at the notch as a function of the number of elements used to



NOTES: 1) $P_1 = \frac{5500}{N}$ WHERE N EQUALS NUMBER OF NODES ON THE INDICATED EDGE

2) BOUNDARY CONDITIONS APPLIED AT EACH NODE ALONG EDGES

3) $E = .30 \times 10^6$ PSI
 $\nu = .3$
 $T = 70^\circ$ F

4) ALGOR DECODER PARAMETERS APPLIED:
 ELEMENTS: TYPE: 2-DIM
 ELEMENTS: INFO: TYPE: PLANE STRESS
 ELEMENTS: GROUP: STEEL MATERIAL PROPERTIES FROM ALGOR'S LIBRARY
 ELEMENTS: COLOR: THICKNES = 1
 TREF = 70

ANALYSIS: STATIC
 GLOBAL: LOAD CASE: A=0, B=0, C=0, D=0
 DECODE: ALL, INTERSECT LINES
 DEFAULT VALUES USED FOR ALL OTHER PARAMETERS

FIGURE 6

Plate w/U-Notch: Stress

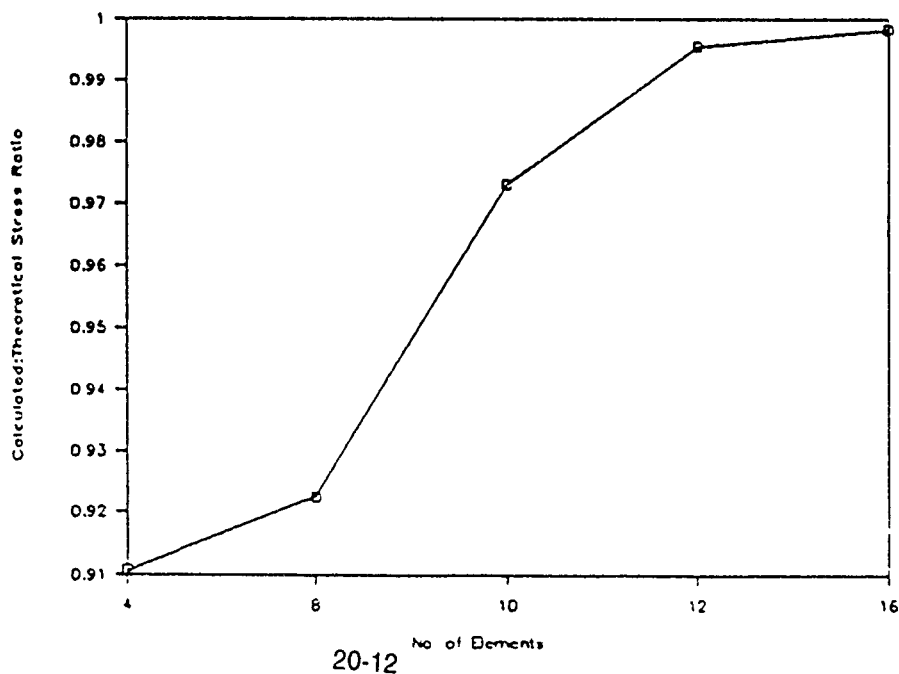
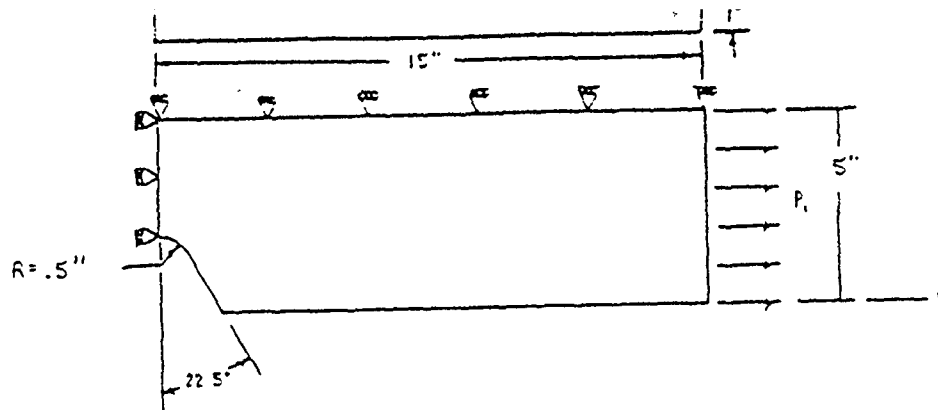
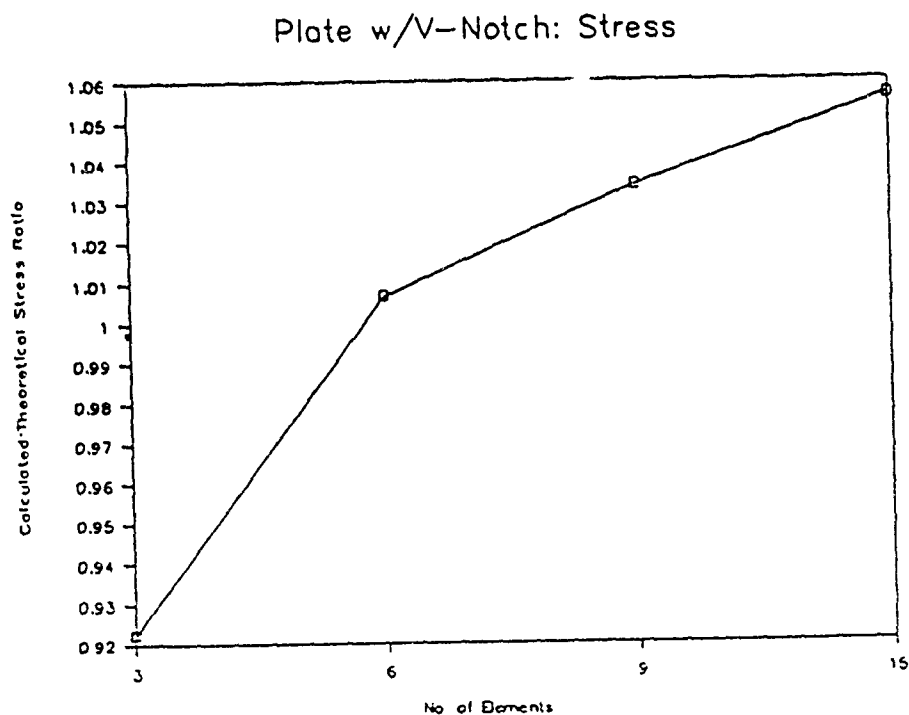


FIGURE 7



- NOTES: 1) $P_i = \frac{5500}{N}$ WHERE N EQUALS NUMBER OF NODES ON THE INDICATED EDGE
- 2) BOUNDARY CONDITIONS APPLIED AT EACH NODE ALONG EDGES
- 3) $E = 30 \times 10^6$ PSI
 $\nu = .3$
 $T = 70^\circ F$
- 4) ALGOR DECODER PARAMETERS APPLIED:
 ELEMENTS: TYPE: 2-DIM
 ELEMENTS: INFO: TYPE: PLANE STRESS
 ELEMENTS: GROUP: STEEL MATERIAL PROPERTIES FROM ALGOR'S LIBRARY
 ELEMENTS: COLOR: THICKNES = 1
TREF = 70
 ANALYSIS: STATIC
 GLOBAL: LOAD CASE: A=0, B=0, C=0, D=0
 DECODE: ALL, INTERSECT LINES
DEFAULT VALUES USED FOR ALL OTHER PARAMETERS

FIGURE 8

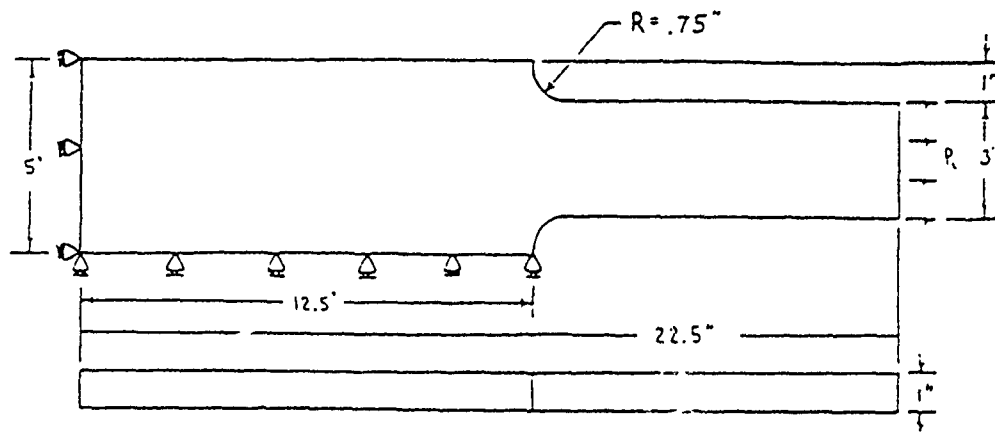


model the curved edge of notch was evaluated and compared to theoretical stress. The plate was modeled using 2-D quadrilateral elements. Axial tension was applied. The notch was divided into 4, 8, 16, and 24 sections. The theoretical stress was calculated using the formula for a square shoulder with fillet in a member of rectangular section found in Reference 6. The stress ratio (calculated/theoretical) as a function of the division of the curved edge of the fillet is shown in Figure 11.

F. Curved Beam with Applied Force on End

A curved beam with a 90 degree arc, clamped fixed end, and free loaded end was modeled using plate elements. See Figure 12. Normally, beam elements would be used. However, this was not feasible since the Algor package used did not support graphical input and output for beam elements. The model using plate elements was deemed significant since analysis of structural features sometimes suggests the use of plate elements in lieu of beam elements.

The program calculated deflection for an applied load was compared to a theoretical deflection calculated by the strain energy method. (See Reference 1.) The curved beam was divided into 2, 4, 6, 8, 9, 12, and 18 sections. Each section was then divided into 5, 6, and 12 elements. The deflection ratio (calculated/theoretical) as a function of the number of sections of the curved beam and the number of elements per section is shown in Figure 13.



Notes: 1) $P_i = \frac{11000}{N}$ WHERE N EQUALS THE NUMBER OF NODES ON THE INDICATED EDGE

2) BOUNDARY CONDITIONS APPLIED AT EACH NODE ALONG EDGES

3) $E = 30 \times 10^6$ PSI
 $\nu = .3$
 $T = 70^\circ$ F

4) ALGOR DECODER PARAMETERS APPLIED:

ELEMENTS: TYPE: 2-DIM

ELEMENTS: INFO: TYPE: PLANE STRESS

ELEMENTS: GROUP: STEEL MATERIAL PROPERTIES FROM ALGOR'S LIBRARY

ELEMENTS: COLOR: THICKNES = 1

TREF = 70

ANALYSIS: STATIC

GLOBAL: LOAD CASE: A=0, B=0, C=0, D=0

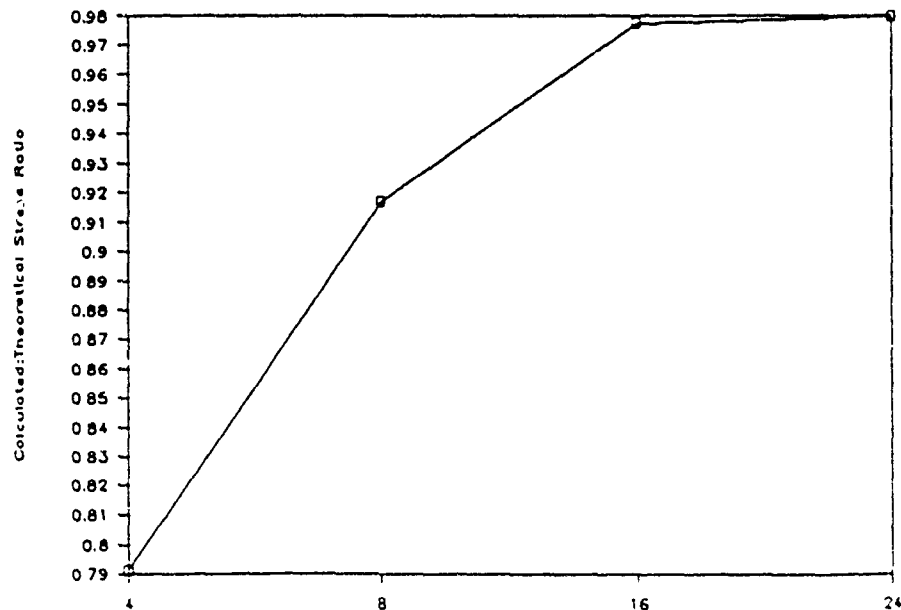
DECODE: ALL, INTERSECT LINES

DEFAULT VALUES USED FOR ALL OTHER PARAMETERS

5) ALGOR HAS A CONSTRUCT: FILLET COMMAND; BUT, IT WAS NOT USED.

FIGURE 10

Plate w/Fillet: Stress



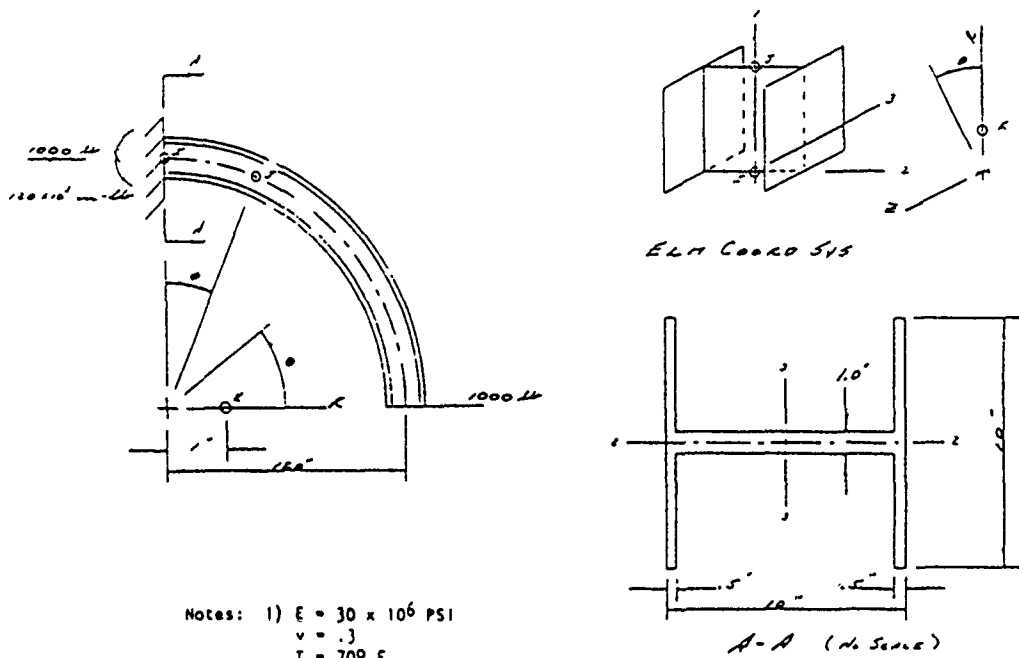
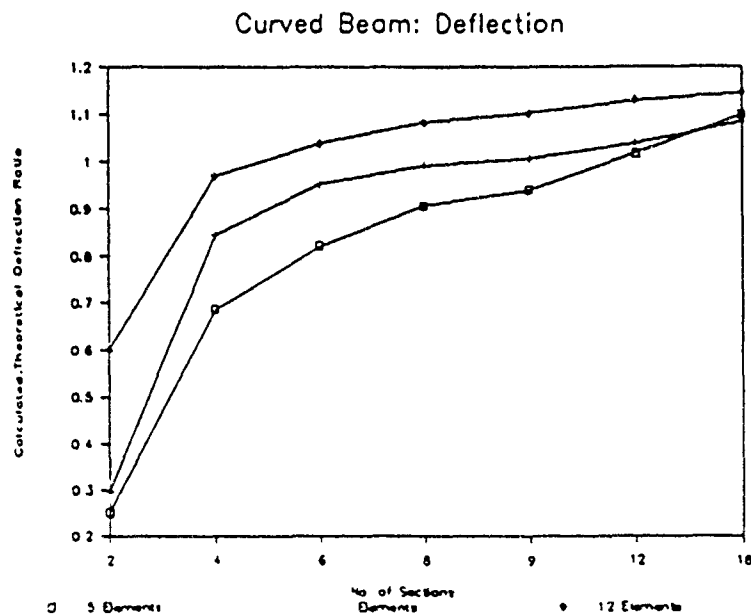


FIGURE 12



VI. RECOMMENDATIONS:

It is recommended that development of the Mesh Density Handbook be continued. While evaluating stress ratios as a function of the number of divisions of the curved edge of the plate models was valid, it was observed that the accuracy of the computer calculated stresses is also a function of the total number of elements in the model. Collection of additional data for the models of the flat plates would improve the accuracy of the generated curves. Also, results from the model of the flat plate with v-notches were not as expected. (The stress ratio exceeded one.) A second model of a flat plate with v-notches would be useful in verifying this data.

REFERENCES

- 1) Bruhn, E. F., Analysis & Design of Flight Vehicle Structures, Cincinnati, OH, Tri-State Offset Company, 1965.
- 2) Fornaro, Gene F., "An Introduction to the Finite Element Method Using ALGOR Supersap," National Computer Graphics Association Conference, Anaheim, CA, March 1990, pp. 169-175.
- 3) Nash, William A., Schaum's Outline Series: Theory and Problems of Strength of Materials, 2nd Ed., USA, McGraw-Hill Book Company, 1989.
- 4) Robinson, John, Ed., "Algor Interactive Systems, Inc.," Finite Element News, No. 1, (February 1990), pp. 5-9.
- 5) Weingarten, Victor, "How Accurate is FEA?," Machine Design, Vol. 61, No. 17, (August 1989), pg. 103.
- 6) Young, Warren C., Roark's Formulas for Stress & Strain, 6th Ed., USA, McGraw-Hill Book Company, 1989.

THE EFFECT OF DOPPLER SHIFT ON THE APPARENT PLUME INFRARED SIGNATURES OF GAS TURBINE ENGINES

Charles R. Hewitt, Jr.
Research Associate # 408

Abstract

Aircraft engine exhaust plumes are important contributors to the total aircraft infrared (IR) signature. Accurate prediction of the plume infrared signature is therefore important in determination of an aircraft's detectability by infrared seekers. Current prediction techniques for plume IR signatures neglect Doppler shifts which arise from the high relative speed of the exhaust gases in relation to the atmosphere. The purpose of this study was to determine what the effect of including the Doppler shift in signature predictions was on the apparent plume IR signature of gas turbine powered aircraft. Results to date indicate that within the range of altitude and velocity at which current aircraft operate, the Doppler shift effects are small for IR seekers working in the 4 to 5 micron waveband. The results do indicate the potential for large effects under special circumstances.

Introduction

Military jet aircraft produce infrared emissions which can be exploited by a weapons system designer for detection, tracking, and guidance. As the emphasis on stealth technology grows, passive detection and tracking by infrared radiation will grow in importance. Already the current generation of Soviet built fighters (e.g. MiG-29) incorporate infrared search and track (IRST) systems which can be used to detect enemy aircraft without alerting the enemy to their own presence. With the growing importance of IRST systems and IR guided missiles, determination of an aircraft's vulnerability relies heavily on the ability to accurately predict the total aircraft IR signature. Emissions from the hot exhaust gases from the aircraft engines play a large role in the total aircraft signature. Thus, it is important to be able to predict the signature from the exhaust gases with a reasonable degree of accuracy.

The purpose of this study is to determine if significant errors occur in predicted apparent plume infrared signatures as a result of not including Doppler shifting in the calculations.

Infrared emission and absorption by gases is very dependent on electromagnetic frequency. A plot of idealized gas emission as a function of frequency can be found in Figure 1a. Narrow bands of emission, called lines, are separated from each other by distances greater than the width of each emission line. The width and the location of the center of each line are determined from quantum mechanical considerations. An idealized absorption plot will have a similar appearance. If one layer of gas is viewed through a second layer (a simple characterization of a plume viewed through the atmosphere), the spectrum of the first will overlay the spectrum of the second if the relative velocity of the two layers is zero. As a result, the energy emitted by the first layer will be rapidly absorbed by the second layer. If, on the other hand, the two layers do have a significant velocity relative to each other (along the direction of propagation), then the first layer's emission spectrum will be Doppler shifted relative to the absorption spectrum of the second layer (refer to Figure 1b). In this case the transmitted energy will be much greater than in the no relative motion case. Calculation of the transmitted energy from an actual plume is complicated by the details of the line shapes and line spacings. Determination of the actual amount of transmitted energy compared to that calculated assuming no relative velocity between the gas layers is the intent of this study.

The scope of this study is limited to the operating regime of gas turbine engines.

Discussion

Infrared radiation in the exhaust plume of jet aircraft comes principally from photons emitted from vibrationally and rotationally excited carbon dioxide (CO₂) and water molecules as they transition from a high energy state to a lower state. A photon's energy is determined by the difference of the energy stored in a molecule's upper state and lower state. The total internal energy of a molecule is composed of the energy stored in various modes of rotation and vibration. CO₂, for example, has three modes of vibration: symmetric stretch, asymmetric stretch, and bending. Each mode has its own set of energies, defined by quantum numbers and quantum selection rules. Photon emission occurs when a mode transitions from one allowed state to the next lower allowed state. Transitions can occur when a single mode changes energy or when multiple modes simultaneously change energy. The photon's electromagnetic frequency is related to its energy by Planck's law:

$$E' - E'' = h\nu_0 \quad (1)$$

A typical missile orIRST may be sensitive to energy over a waveband which contains many (upwards of 10,000) transitions or spectral lines.

The photons emitted by a large number of molecules transitioning between the same upper and lower states do not all have the same frequency as implied by equation 1. Instead, there is a distribution of frequencies, centered on the frequency given by (1). The spread of frequencies, known as line broadening, occurs for several reasons. Doppler, Lorentz, and Voigt broadening are the types of broadening of most significance to gas turbine infrared signatures. Doppler broadening is the result of relative motion between the emitter and the observer which causes the observed frequency to shift either up or down in the spectrum depending on whether the emitter is moving toward or away from the observer. The magnitude of the shift depends on the relative velocity and the emitted frequency, or

$$\nu' = \nu_0 (1 \pm V/c) \quad (2)$$

A gas at rest is composed of a large number of molecules moving at random velocities. This random motion gives rise to what is macroscopically known as temperature. The Maxwell-Boltzmann distribution gives the probability distribution of molecular velocities for a gas in thermodynamic equilibrium as a function of the gas's temperature:

$$p(V) = (m/2\pi KT)^{1/2} \exp(-mV^2/2KT) \quad (3)$$

When the Doppler shift is applied to the molecular velocity distribution, the result is a Doppler broadened probability of emission (or absorption) given by

$$p(\nu) = (1/\gamma_D \pi) \exp\{-(\nu - \nu_0)^2 / \gamma_D^2\} \quad (4)$$

where

$$\gamma_D = (\nu_0/c) (2KT/m)^{1/2} \quad (5)$$

is the Doppler width (the spectral distance from line center to the point where the probability is 50% of the probability at line center is $\gamma_D \ln 2$).

Lorentz broadening results from molecular collisions. The Doppler broadening described above was assumed to result from the spontaneous emission of photons from molecules uninfluenced by other molecules. In gases under pressures typical of the atmosphere, the molecular number density is large enough that collisions between molecules occur at a fairly high rate. A collision may cause a transition to occur in one or both of the molecules. The

effect of a collision induced transition is that the wave-train of the photon is disrupted to the extent that it can no longer be represented by a single frequency. Instead, a set of frequencies is required. When a large group of molecules is considered, a distribution of frequencies results which is given by

$$p(\nu) = (\gamma_L/\pi)/[(\nu-\nu_0)^2 + \gamma_L^2] \quad (6)$$

where γ_L is the Lorentz half-width. The Lorentz half-width is effected by the gases which are colliding with the emitting molecule. An empirical form has been established for the CO₂ half-width when considering the mixture of gases in the atmosphere¹ as

$$\gamma_L = 0.07 P (296/T)^{0.7} \quad (7)$$

While the mixture of gases in the exhaust of a gas turbine differs from the atmosphere, the dominating gas continues to be nitrogen. Therefore, the same form of Lorentz half-width can be used for the exhaust gases as for the atmosphere to a reasonable degree of accuracy.

Voigt broadening is the result of Doppler and Lorentz broadening occurring simultaneously. The spectral profile under these circumstances, known as the Voigt profile, is a convolution of the Doppler and Lorentz profiles and is given by

$$p(\nu) = (1/\gamma_D \pi^{1/2}) (a/\pi) \int_{-\infty}^{\infty} \{ \exp(-y^2)/[a^2 + (\xi-y)^2] \} dy \quad (8)$$

where

$$a = \gamma_L/\gamma_D$$

and

$$\xi = (\nu - \nu_0)/\gamma_D$$

The half-width at half height for a Voigt profile has no closed form expression. An approximate method for determining the Voigt half-width is given by Olivero².

Figure 2 shows the line half-widths as a function of altitude for atmospheric CO₂. Atmospheric pressures and temperatures for Figure 2 were taken from the 1962 U.S. Standard Atmosphere³. The Doppler half-width is based on a wavenumber of 2250 cm⁻¹.

Variation in the temperature dependent Doppler half-width is considerably smaller than for the pressure dependent Lorentz half-width due to the fact that altitude variations in temperature are much smaller than in pressure. Note that the Voigt half-width approaches the Doppler half-width when γ_D is larger than γ_L . Likewise, γ_V approaches γ_L when Lorentz broadening predominates.

The absorption coefficient is the fractional change in energy per unit optical depth. The spectral absorption coefficient is related to the spectral line profiles given in (4), (6), and (8) by the line strength:

$$k(\nu) = p(\nu) S \quad (9)$$

The line strength S accounts for the probability that a particular transition can occur and for the population of the initial state (i.e. upper state in emission; lower state in absorption). Note that the absorption coefficient could alternatively be written as

$$k(z) = C L(z) \quad (10)$$

where C is a grouping of constants, $L(z)$ is a line shape function which depends on the type of line, and z is the normalized spectral distance or

$$z = (\nu - \nu_0)/\gamma \quad (11)$$

The spectral intensity reaching a viewer from a layer of emitting gas (e.g. aircraft plume) with a set of properties J propagating through a layer of absorbing gas with properties A (refer to Figure 3) is given by

$$I_\nu = \tau_{\nu,A} \epsilon_{\nu,J} I_{bb}(T_J) \quad (12)$$

where the transmissivity of layer A , $\tau_{\nu,A}$, is

$$\tau_{\nu,A} = \exp(-k_{\nu,A} x_A) \quad (13)$$

and the emissivity of layer J , $\epsilon_{\nu,J}$, is

$$\epsilon_{\nu,J} = 1 - \exp(-k_{\nu,J} x_J) \quad (14)$$

The optical depth (i.e. the product of absorption coefficient and path length) of each layer plays an important role in determining the intensity emitted from or transmitted through a layer. At small optical depths, the profile of emissivity or transmissivity has the same shape as the absorption coefficient profile. At large optical depths, the spectral profile of ϵ or τ broadens due to the character of the exponential function. At some optical depth, the line center (i.e. when $\nu = \nu_0$) saturates, making the emissivity virtually 1.0 (or the transmissivity 0.0). As the optical depth increases from this point, regions just above and below the line center saturate also. The effect is to give the line a broader profile than would be expected from the absorption coefficient profile. Figure 4 shows an example of the effective broadening for the emissivity, normalized to the line center emissivity, of a layer which exhibits Lorentz broadening. The spectral width over which the emissivity is greater than 50% of the line center emissivity greatly increases as the optical depth becomes large.

By defining the effective half-width, $z_{1/2}$, as the distance from line center, measured in absorption coefficient profile half-widths, to the point at which the emissivity profile is 50% of the maximum (i.e. line center) value, the value of $z_{1/2}$ can be found from equations (4), (6) and (14). For a Doppler broadened profile, the effective half-width is

$$z_{1/2} = \{-\ln [\ln(0.5 + 0.5\exp(-Q))/(-Q)]\}^{1/2} \quad (15)$$

and for a Lorentz broadened profile

$$z_{1/2} = \{-1 - Q/\ln[0.5 + 0.5\exp(-Q)]\}^{1/2} \quad (16)$$

where Q is the product of the grouped constant in (10) and the path length, x .

Effective half-widths come into play when determining if a line is Doppler shifted enough to cause a significant difference in the transmitted energy. As discussed above, emitter motion at some velocity relative to an observer causes the observed frequency to be shifted from the emitted frequency. The Doppler broadened profile accounts for this phenomenon for a collection of molecules whose average velocity is zero (i.e. the collection has a thermal velocity distribution). In the exhaust of gas turbines, significant non-zero bulk gas velocities exist. When the angle between the jet plume axis and the viewing direction is considered, the observer relative velocity ranges from zero (side view position) to \pm the net gas velocity (the jet velocity calculated from compressible flow relations minus the aircraft flight velocity) for

tail-on or nose-on views. Thus, an extra Doppler shift is superimposed on the jet emission profile due to the en masse motion of the emitting molecules found in the jet plume relative to the atmospheric absorbing molecules. The extra Doppler shift is superimposed regardless of the type of absorption coefficient profile which defines the local line structure. Whether or not the Doppler shift is sufficient to cause a difference in the apparent signature depends on the effective half-width of both the absorbing and emitting layers.

An example of the spectral apparent intensity due to a single isolated line for a range of Doppler shifts can be found in Figure 5 for the geometry shown in Figure 3. A curve for the source intensity has also been plotted for reference. The amount of shift for each line in Figure 5 is measured in terms of the jet half-width. Likewise, the horizontal axis is measured in terms of the jet line half-width. Comparing the curve for a shift of 0.5 line half-widths to the no shift curve indicates that, when integrated across the spectrum shown, a slightly greater total intensity could be expected. As the shift increases, the spectral apparent intensity takes on a shape similar to the curve for the unshifted source value. Eventually, at large enough shifts, the apparent spectral intensity acts independently of the transmission profile. The result is a large difference in the total transmitted energy compared to the unshifted total.

Figure 6 is a plot of the integrated apparent intensity for situations as described in Figure 5. At small atmosphere optical depths, the total transmitted intensity is essentially independent of the shift imposed on the line profile. As the atmospheric optical depth increases, the change in transmitted intensity for a given shift increases. The shift required to have the apparent intensity approach the source intensity also increases with optical depth. If the shift is large enough the apparent intensity is independent of atmosphere optical depth as was indicated in Figure 5. The upshot of Figure 6 is that for a single line, the total apparent intensity can be seriously underestimated if the line Doppler shift is not included in the calculation. Calculations at other conditions representative of those found in the 4.3 micron CO₂ band and military aircraft flight conditions show essentially the same trends but vary somewhat in their particulars.

Given that shifts of sufficient magnitude can cause significant errors in the total intensity for a single line, two questions need to be answered:

- Are the velocities which can be found in jet aircraft large enough to cause significant Doppler shifts?

- Can the single line results be used to extrapolate to wavebands with many lines?

Typical plume velocities were determined using a thermodynamic cycle analysis technique developed by Mattingly⁴. An assumption was made that a turbojet cycle operating without afterburning would be representative of flight points of interest. Further, the engine cycle was set to have an overall pressure ratio of 16. The exit velocities for a range of aircraft flight altitudes and Mach numbers, assuming a fixed, maximum power setting were calculated using the cycle analysis and one dimensional isentropic flow relations. The maximum relative velocity, i.e.

$$V_{rel} = V_{jet} - V_{aircraft} \quad (17)$$

was found for each flight condition. The results are plotted in Figure 7. The variation in relative velocity with flight altitude is small. The relative velocity is, however, inversely proportional to the flight Mach number. Taking 830 m/s as a representative velocity and 2250 cm^{-1} as the wavenumber, it is found using (2) that Doppler shifts on the order of $6 \times 10^{-3} \text{ cm}^{-1}$ can be expected in the emission spectrum of gas turbine plumes. Referring to Figure 2, it can be seen that the Voigt half-width is less than or equal to $6 \times 10^{-3} \text{ cm}^{-1}$ at altitudes of 60,000 feet or greater. The magnitude of the Doppler shift in relation to the Voigt half-widths and the low flight Mach number result from above indicate that the most likely scenario for there to be significant error in predicted plume signatures due to not accounting for the Doppler shift is in the upper left hand corner of the flight envelope at high throttle settings.

The ability to extrapolate the single line Doppler shift results to the expected signature for wide spectral bands depends on the line spacing, or the inverse of line spacing (which is the number of lines per unit wavenumber). If two absorption lines are so close that when an emission line shifts it moves from one absorption line to another, the total apparent energy will not be appreciably different than if the emission line had never been shifted. A complication here is that the line spacing is a strong function of temperature. Figure 8 shows the line spacing, averaged over a 5 cm^{-1} window, as a function of wavenumber for a series of gas temperatures. The curve for 200 K is in the realm of atmospheric temperatures for 30,000 to 100,000 feet. Exhaust gas static temperatures for the engine cycle described above are approximately 600 K. A third curve at 400 K has been shown in order to give some idea of

the functional dependence of line spacing on temperature. Figure 8 shows that in the atmosphere (i.e. where the absorption takes place), the line spacing is on the order of 10^{-1} cm^{-1} or greater. Previously, it was shown that the atmospheric Voigt half-widths ranged from 10^{-1} cm^{-1} at the lower altitudes to 10^{-3} cm^{-1} at the upper altitudes. Taken together, it can be seen that the lines are too closely spaced in the lower atmosphere for the Doppler shifting of emission lines to cause any great effect. At altitudes of 60,000 feet and greater, the atmospheric absorption lines begin to separate enough that Doppler shift effects might begin to be noticed.

The inverse of the line spacing, or the number of lines per unit wavenumber effects the ability to extrapolate single line results to wide band results also. The number of lines present in a layer is dependent on the number of energy states which have significant populations. As temperature increases, the higher energy state populations increase at the expense of the lower energy state populations. The result is that hot gas layers, similar to jet plumes, have many more emission lines than cool layers, like the atmosphere have absorption lines. Figure 9 shows the number of lines located within a 5 cm^{-1} waveband (e.g. 2225 to 2230 cm^{-1}) as a function of temperature for a select set of spectral intervals. The ratio of lines for a layer at 600 K to lines for a layer at 200 K is 15 to 20 for the spectral intervals shown. Thus, for every absorbing line in the atmosphere, there are 15 or so emitting lines in the plume. The transmitted energy for the extra lines is independent of the Doppler shift since there is no corresponding absorption line. Therefore, the energy recovered from the few lines where there is correlation between the plume and the atmosphere is a small percentage of the total energy. The line density ratio (for 600 K to 200 K) varies with wavenumber from around 4 to as much as 90 for the 4.3 micron band of CO_2 . Since the line density ratio varies so much, extrapolating from single line to broadband apparent intensities directly will be prone to error. On the other hand, taking the intermediate step of extrapolating the single line results to a 5 cm^{-1} band and then summing the apparent band energies across the wide band is likely to give respectable results.

Results

Using the technique suggested above one could estimate the discrepancy in the transmitted radiance for a particular scenario. For example, the turbojet engine cycle discussed above gave a relative velocity of 830 m/s. The Doppler shift for this velocity at 2240 cm^{-1} is $6.197 \times 10^{-2} \text{ cm}^{-1}$. The atmospheric Voigt half-width, assuming the aircraft is operating at

100,000 feet, is $2.364 \times 10^{-2} \text{ cm}^{-1}$. Likewise, the jet Voigt half-width is $3.318 \times 10^{-2} \text{ cm}^{-1}$. In both cases the broadening is principally Doppler. Using Figure 6, which was calculated for the conditions at 100,000 feet and the ratio of the Doppler shift to the jet half-width of 0.7, the proportion of energy which is transmitted accounting for the Doppler shift is 0.90909. Likewise, the transmitted energy for no shift is .23864. Now, assuming that 600 K is the jet static temperature and 220 K is the atmosphere temperature, Figure 8 can be used to determine the line density ratio for the $2240\text{-}2245 \text{ cm}^{-1}$ band as 16.8. Then the error incurred by not including Doppler shift effects in the signature predictions is found by

$$\begin{aligned} \% \text{ Error} &= \frac{(16.8 - 1) + .90909}{(16.8 - 1) + .23864} - 1 \\ &= 4.2 \% \end{aligned}$$

A similar procedure for the $2245\text{-}2250 \text{ cm}^{-1}$ waveband yields an error of 5.1 %. Based on these results, an error for the entire 2240-2250 waveband, should be on the order of 5 %.

A separate line by line (LBL) calculation shows an error of 5.3% for the same conditions. A cursory examination of the line density plots indicates that for most of the 5 cm^{-1} increments in the 4.3 micron CO_2 band, the line density ratio is as large or larger than the line density ratio at 2240 or 2245 cm^{-1} . Thus, under the given conditions, an error of 5 % in the predicted broadband signature is an upper bound on the real error.

Conclusions and Recommendations

The purpose of this study was to determine the effect of including line Doppler shifting in calculations of apparent signature. The scope of the study was limited to the emission and absorption of CO_2 in the 4.3 micron band for gas turbine plumes. The results indicate that only small errors (i.e. 5 % or less) will be found in the band radiance. The maximum errors will also occur only in the extreme upper left hand corner of the flight envelope (i.e. very high altitude and low Mach numbers) at nonafterburning throttle settings. A further constraint on these results is that the observer is at the same altitude as the emitting aircraft. Errors considerably larger than 5 % can occur in the spectral details. As a result it is not clear that ignoring the Doppler shift in plume signature calculations is the right thing to do. Instead, using the approach outlined above, an estimate of the error in a given situation could

be used to decide if the extra computational effort is required.

Due to the limited scope of the present work, it is recommended that the investigation be continued by

- Examination of the effect of several layers of plume gases on the apparent signature. This would be more representative of real plume flowfields. The current work provides little insight into what the result of multiple layers will be.
- Extension of the throttle setting range to include afterburning. The larger temperature, and CO₂ concentration both act in a nonlinear fashion on plumes, so the result of accounting for Doppler shifts may contain some surprises.
- Consider the effect of the detector being located at a different altitude (specifically orbital altitudes).

Acknowledgements

This work has been funded by the Air Force Office of Scientific Research under the Summer Research Program for graduate students. In addition, various Sverdrup Technology, Inc./AEDC Group personnel have been helpful in many ways. I would specifically like to acknowledge the help of H. T. Bentley III, J. A. Drakes, and C. C. Limbaugh. Finally, I would like to thank my wife, Elizabeth for putting up with my extended absence from home over the course of the summer.

References

1. H.J.P. Smith et al, "FASCODE - Fast Atmospheric Signature Code (Spectral Transmittance and Radiance)", AFGL-TR-78-0081, 16 January, 1978.
2. J.J. Olivero and R.L. Longbothum, "Empirical Fits to the Voigt Line Width: A Brief Review", JQSRT, v.17, pp.233-236, 1977.
3. Anonymous, "Aeronautical Vestpocket Handbook, 20th Edition", United Technologies, Pratt & Whitney, 1986.
4. J.D. Mattingly, "On-Design and Off-Design Cycle Analysis Computer Programs", American Institute of Aeronautics and Astronautics, January, 1987.

Nomenclature

- c - speed of light
- C - grouping of constants in the definition of absorption coefficient
- E' - energy of the molecule's upper state
- E'' - energy of the molecule's lower state
- h - Planck's constant
- I - intensity of radiation
- I_{bb} - intensity of radiation of a blackbody
- k - absorption coefficient
- K - Boltzmann's constant
- L - line shape function
- m - molecular mass
- $p()$ - probability distribution function
- P - pressure
- S - line strength
- T - temperature
- V - velocity
- x - distance along path
- z - normalized distance from line center

- γ_D - Doppler line width
- γ_L - Lorentz line half-width
- γ_V - Voigt line half-width
- ϵ - emissivity
- ν_0 - photon frequency; line center wavenumber
- ν' - Doppler shifted wavenumber
- τ - transmissivity

Figure 1 (A) Idealized Spectral Emission



Figure 1 (B) Idealized Doppler Shifted Spectral Emission and Transmission

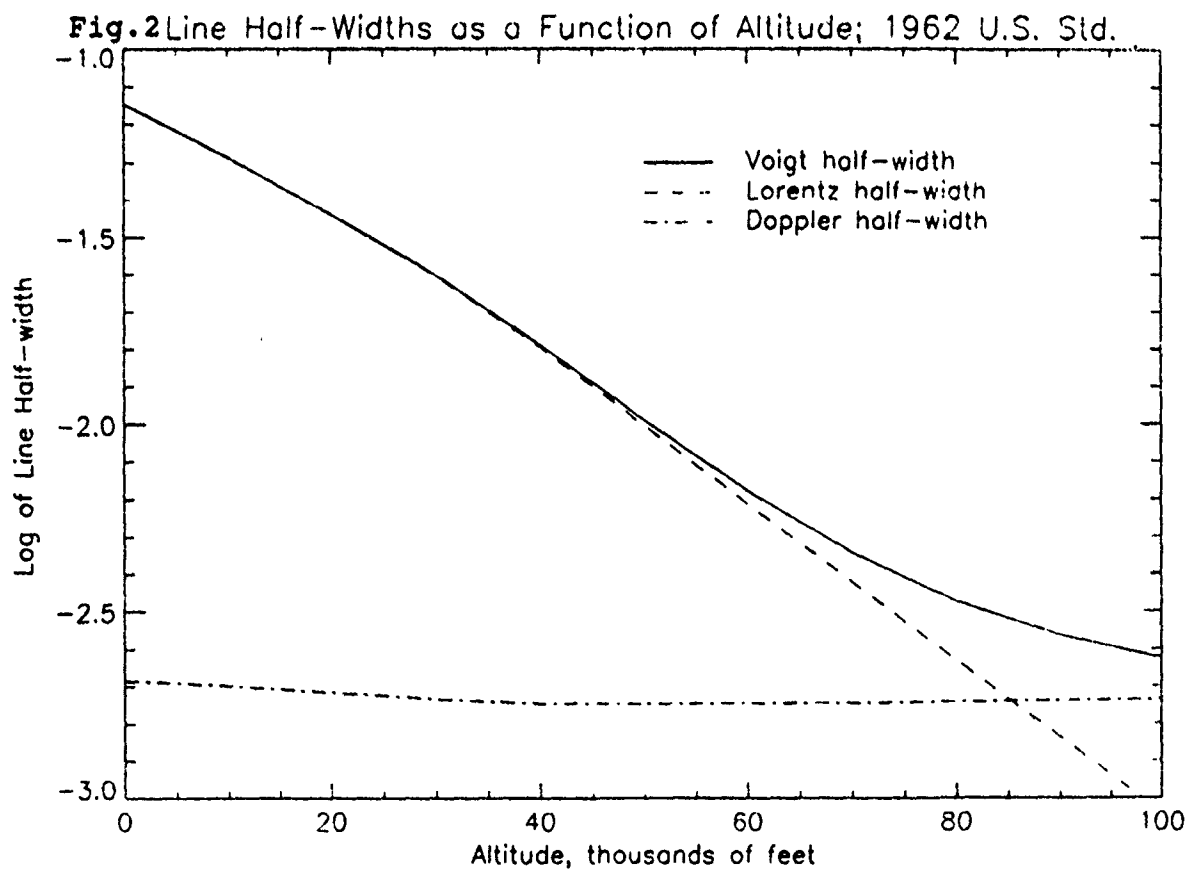
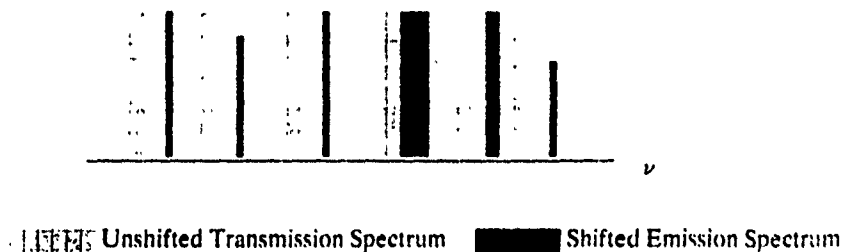


Figure 3. Schematic of the Two Layer Emission-Absorption Problem

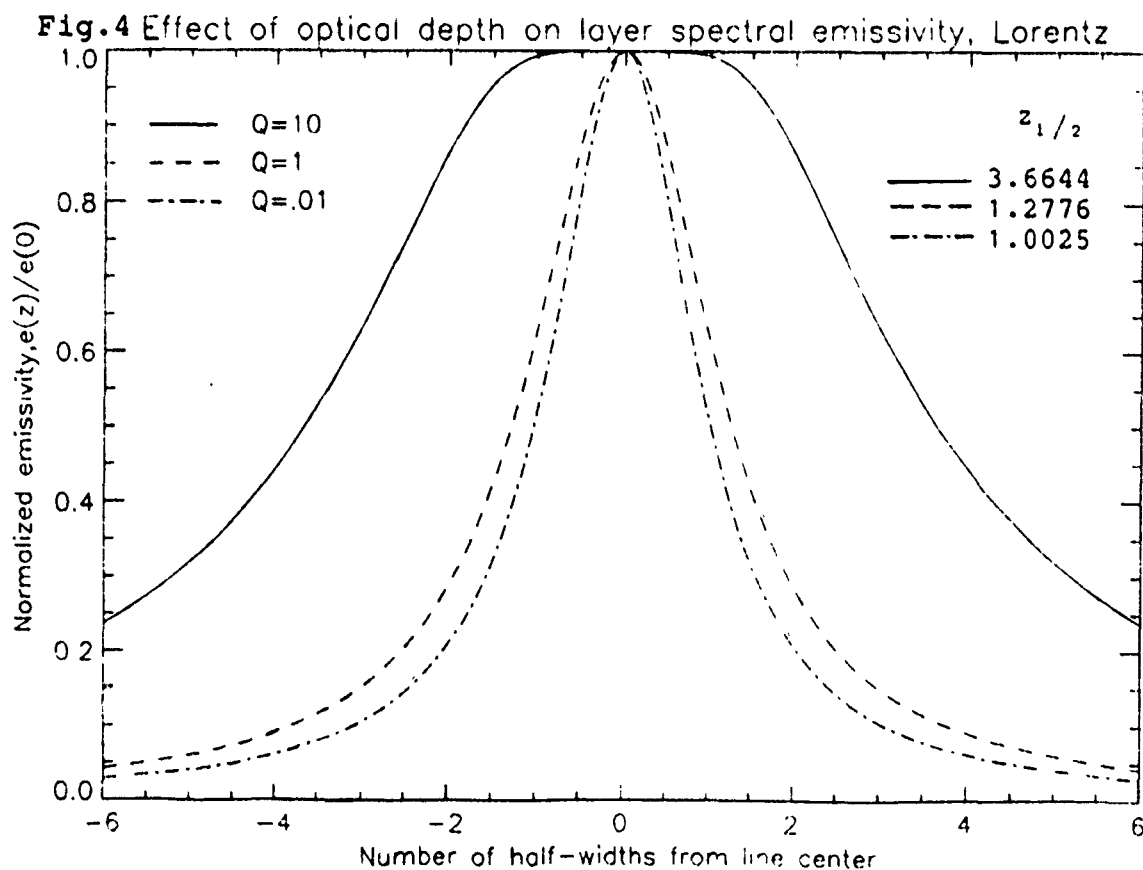
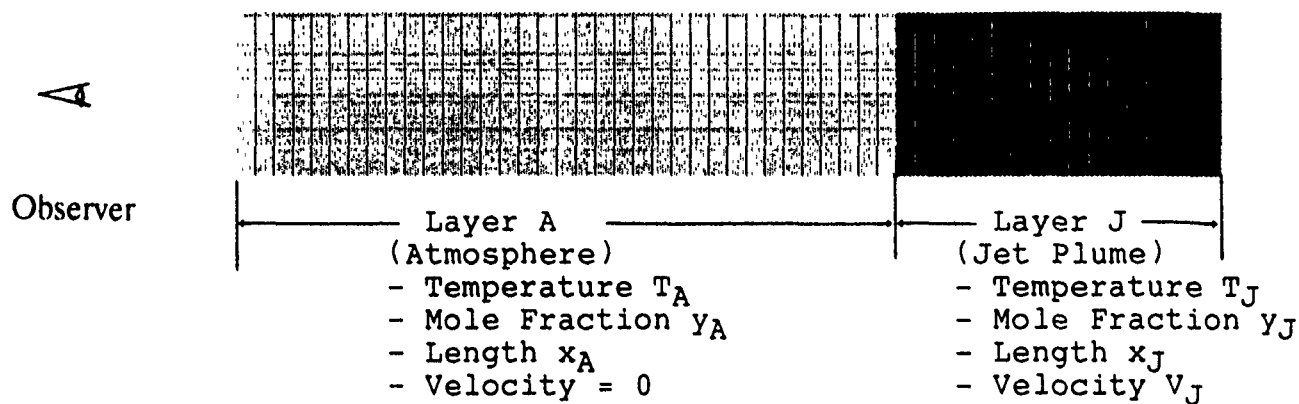


Fig.5 Effect of shift at 100km range

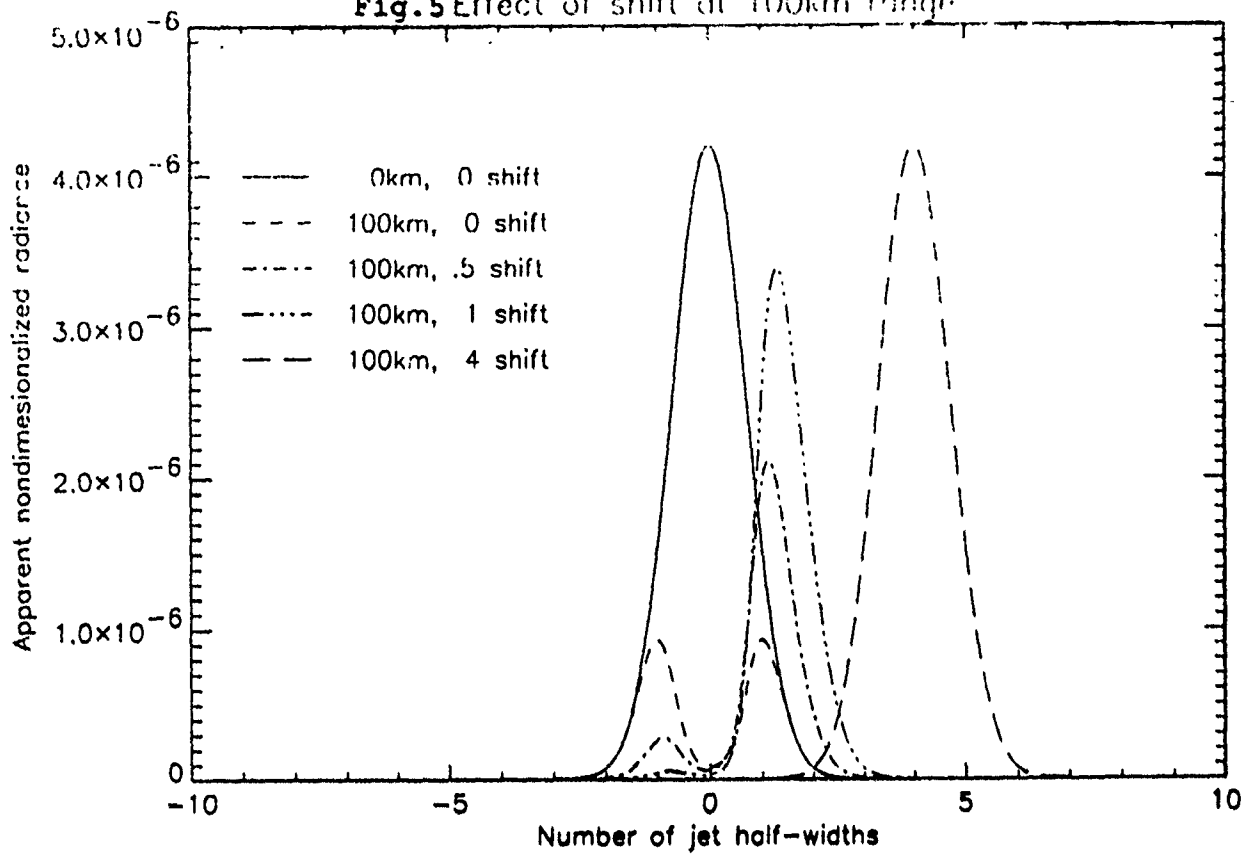
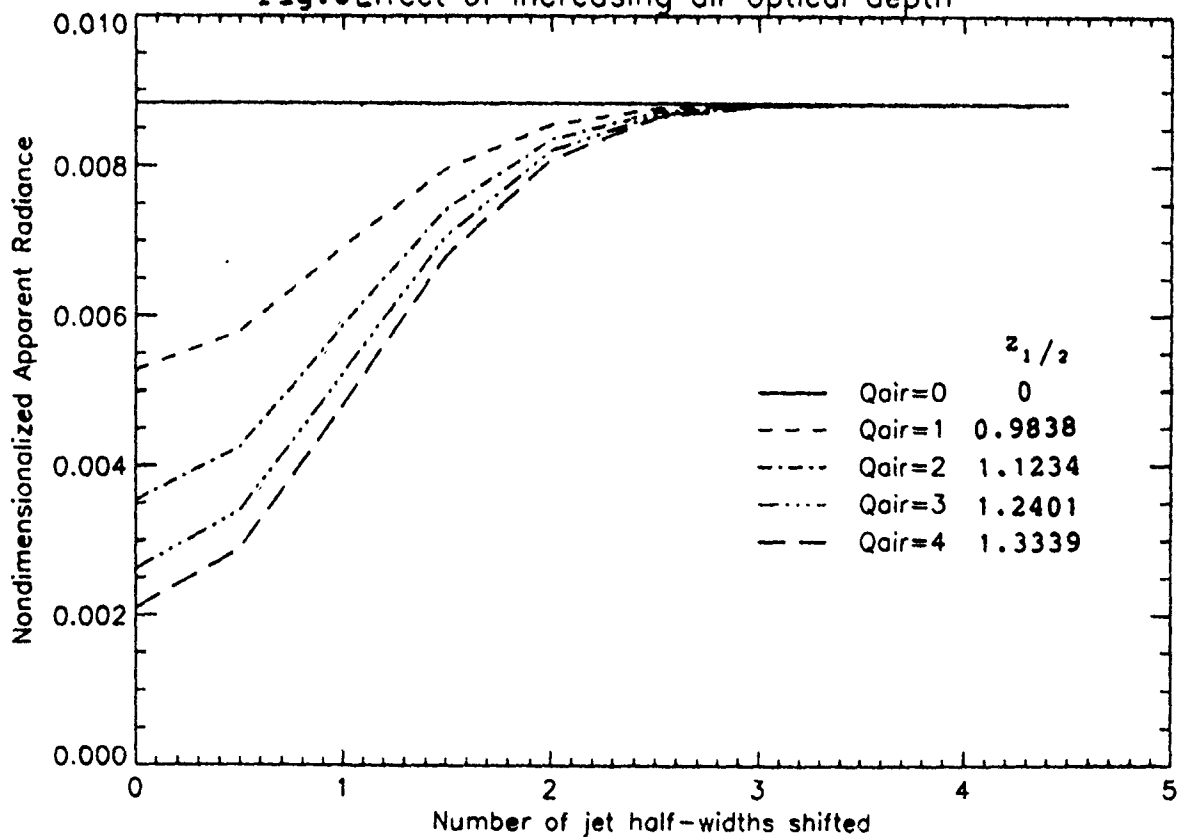


Fig.6 Effect of increasing air optical depth



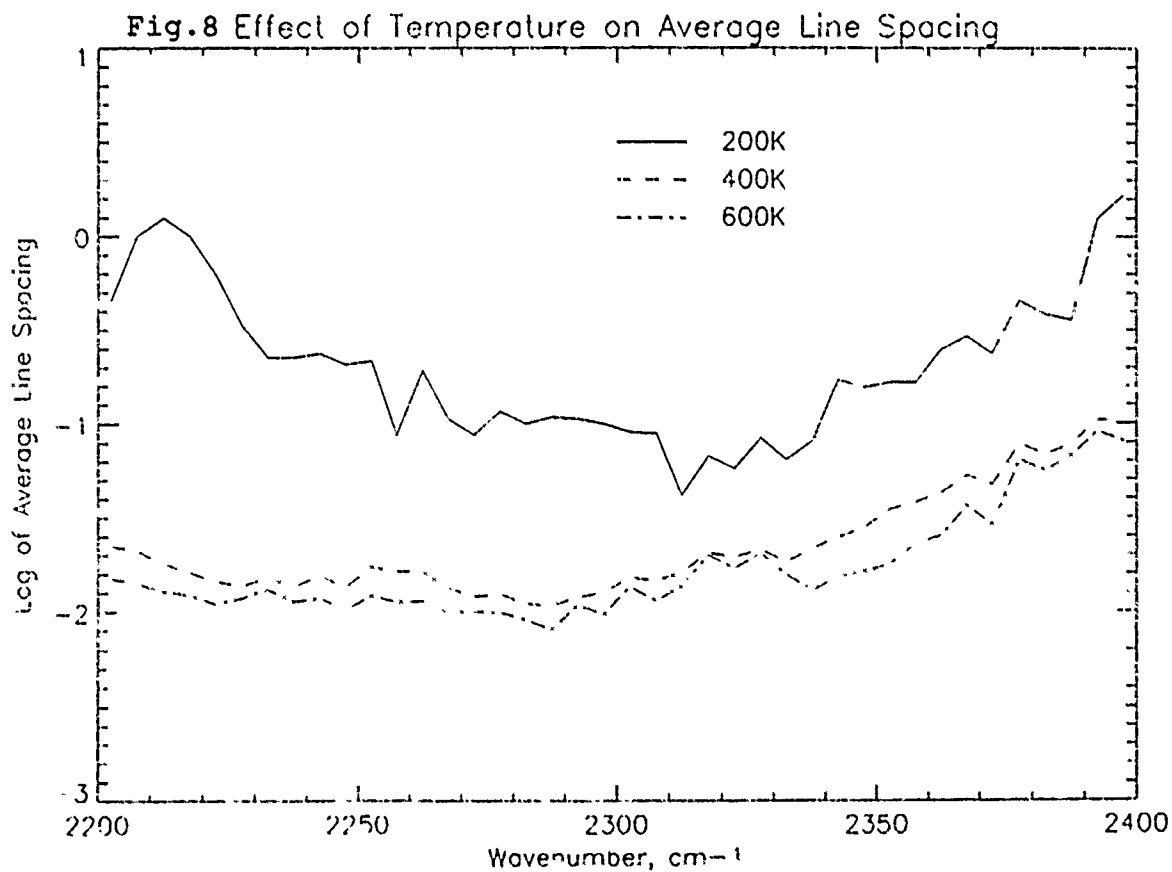
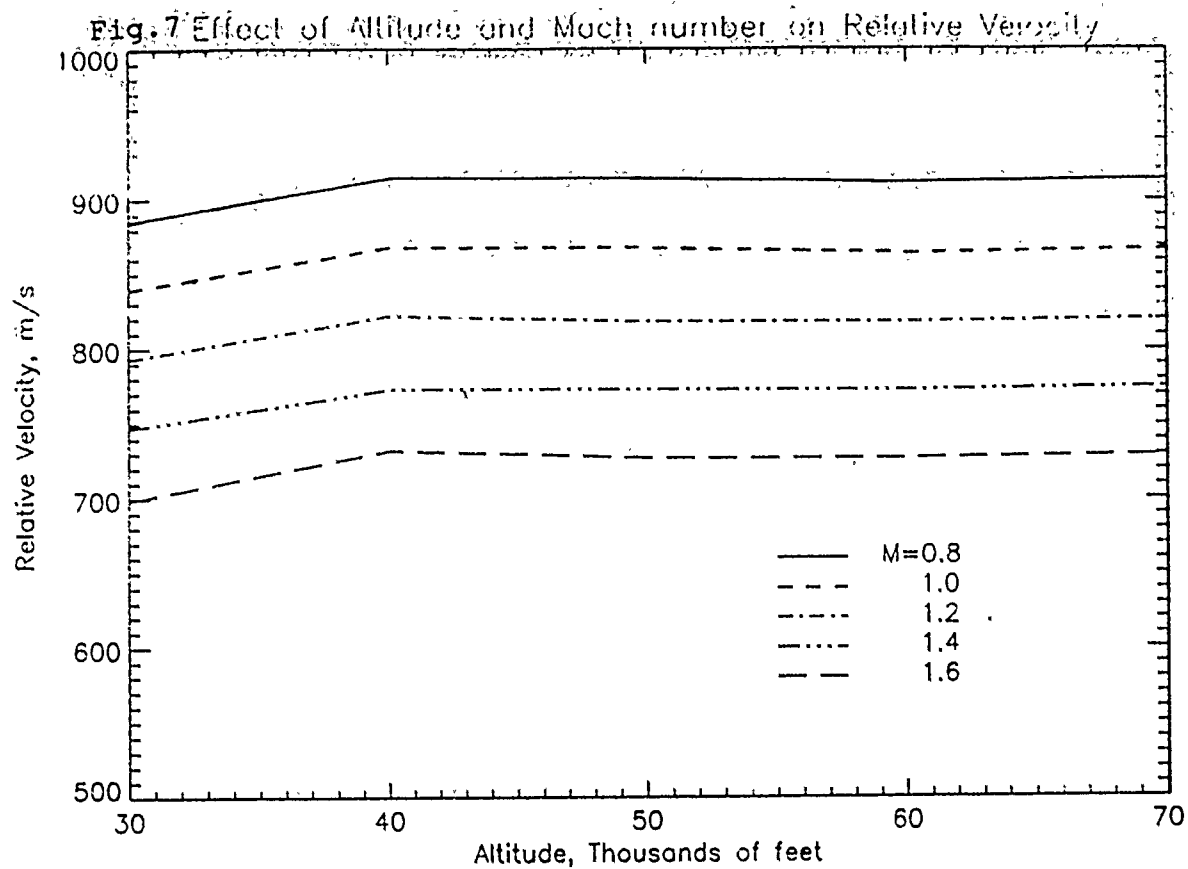
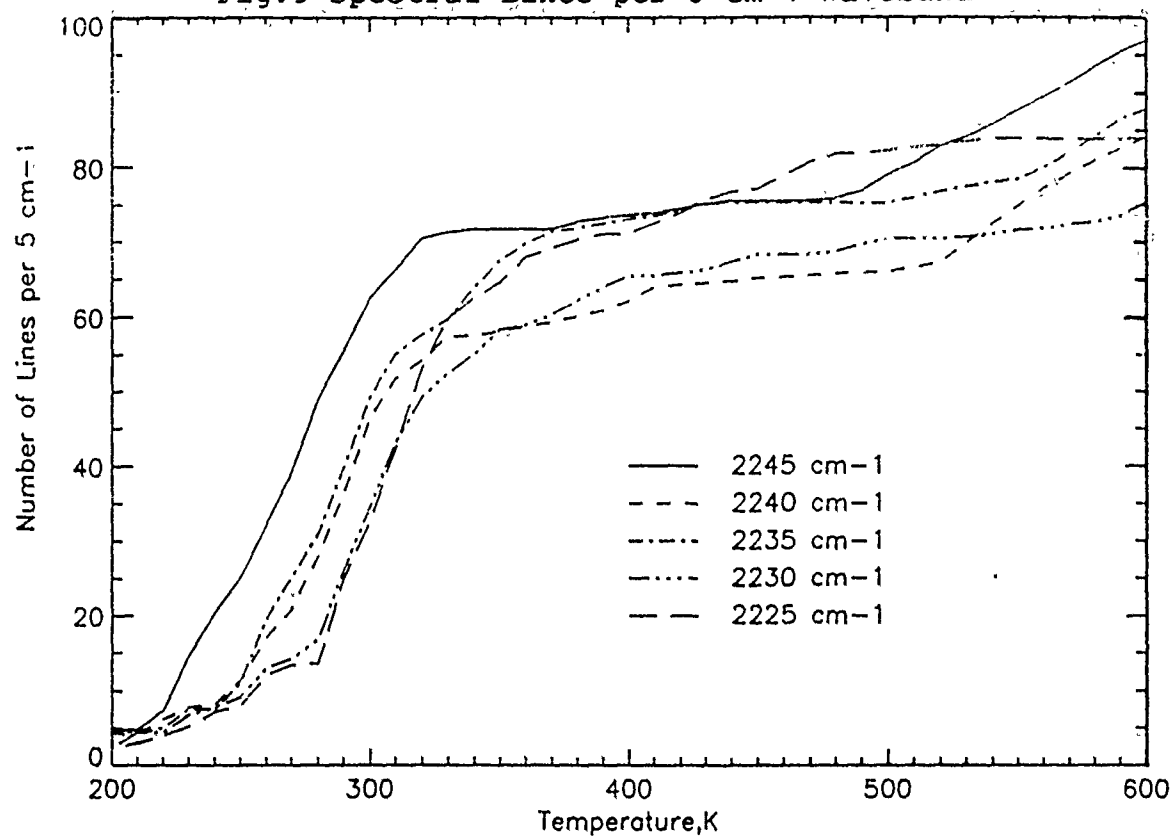


Fig.9 Spectral Lines per 5 cm⁻¹ Waveband



Recommendations Report on Water Tunnel Analysis
of ASTF C1 Plenum Section

by
Stephen D. Howard
AFOSR Summer Research Student
August 13, 1991

Acknowledgments

I would like to thank AFOSR and RDL for giving me the opportunity to do this research. This summer program has also allowed me to gain some excellent work experience. Special thanks to Terri Brock of AEDC for the invaluable assistance she provided throughout the program.

RECOMMENDATIONS REPORT ON WATER TUNNEL ANALYSIS OF ASTF C1 PLENUM SECTION

Stephen D. Howard
AFOSR Summer Research Student

Abstract

The basics of water tunnels are examined in terms of theory and design. The practicality of using a water tunnel model to determine the inlet/plenum design of the C1 test cell of ASTF is examined. A recommendation is made on using a water tunnel approach to the design.

Introduction

The purpose behind this report was to do a feasibility study on whether or not a water tunnel would be effective in analyzing the flow quality of several different plenum concepts. These plenum concepts are being evaluated as part of the C1 test cell in the Aeropropulsion Systems Test Facility (ASTF). ASTF is being modified to support testing of the National Aerospace Plane (NASP). One of the criteria for the plenum is to deliver the highest level of flow quality possible to the inlet of the downstream nozzle. Here, flow quality encompasses uniformity, low turbulence levels, and minimal pressure and temperature range deviations.

First, the basics on water tunnels will be presented for background purposes. Then its utility in the particular case of ASTF will be examined.

Background

1. Water Tunnel Theory

While water tunnels are often used to study hydrodynamic phenomena, they can be used to emulate situations where air is the flow medium. The fact that water is a liquid and air is a gas requires certain considerations to be

made. One of the key differences between the two is their Reynolds number (Re) values. The Reynolds number is defined as,

$$Re = Vx/\nu$$

where,

V = velocity

x = distance

ν = kinematic viscosity

At standard temperatures, ν_{air} is about 15 times greater than ν_{water} . Thus, for similar velocities and lengths, the Reynolds number for water will be approximately 15 times greater than that for air (Ref.1). Unfortunately, many interesting aerodynamic phenomena occur at high air velocities (turbulent flows). It is physically impossible to match these speeds in a water tunnel (flows are essentially laminar). In fact, water tunnel models correctly simulate only scaled down versions of wind tunnels.

There have been attempts to increase the accuracy of the hydrodynamic model. One technique involves raising the water temperature. The kinematic viscosity of water will decrease as the temperature increases (Fig. 1). This will then give a higher value for the Reynolds number. Increasing the Reynolds number through this method allows the water tunnel to be smaller in size. A technical report by NASA (then NACA) demonstrated how a temperature increase could cut the width of a water channel by over 50% (Ref. 2).

Another critical factor which must be considered is what effect compressibility may have on the flow model. Water, being a liquid, has a much lower compressibility factor than air (i.e., water is essentially incompressible). This makes the modelling of compressible flow situations very difficult. Flows involving shocks (supersonic flows) are certainly compressible, and cannot be simulated in a water tunnel. However, there is some evidence which demonstrates some compressible flows can be modelled to a certain extent. With respect to the particular case of external separated flows, it appears that the qualitative aspects remain unchanged when compressibility is introduced.

The time dependence of flows is also a problem confronting potential water tunnel users. Dyes used to visualize the water flow will not accurately

imitate streamlines in unsteady situations. This can pose problems when flow separation occurs. David McDaniel's thesis (Ref. 3) describes this problem in detail.

2. Water Tunnel Design

Water tunnel designs have a resemblance to wind tunnel designs. Normally, water tunnels are closed-circuit and smaller in scale to most wind tunnels. They can range in length and diameter; the Garfield Water Tunnel facilities at Penn State (Ref. 4) have a large water tunnel (100 ft. in length, 12 ft. diameter test section) and a small water tunnel (36 ft. in length, 1 ft. diameter test section).

In addition, water tunnels can be either vertical or horizontal in design. Vertical tunnels normally occupy a smaller amount of space than horizontal ones. Figures 2 and 3 show examples of vertical and horizontal return water channels.

A typical water tunnel design is the UTSI facility (Fig. 4). It is a closed-circuit, horizontal return channel with a 1.5 ft² rectangular test section (Ref. 5). The test section walls are made of plexiglass so that the flows can be photographed. The water is pumped through the tunnel with a double-bladed propeller. The screens and honeycomb meshes located near the entrance of the stilling chamber help reduce the turbulence to acceptable levels ($= 0.1\%$). The ductwork of the UTSI water tunnel is axisymmetric; the transition to the rectangular test section is accomplished with a bell mouth nozzle. Data measurements around the flow model can be processed on the facility's data acquisition system. Figure 5 shows the visualization of strake vortex flow on an F-16 model in the UTSI tunnel.

3. Visualization in the Water Tunnel

A key feature of the water tunnel is its ability to make flows visible. Basically, this is achieved by injecting into the water some sort of substance that can be illuminated. This substance is usually color dyes; but it can be tracer particles, hydrogen bubbles, or other elements. The injection of dye can take place upstream of the flow area of interest using a probe device, or it

can occur at ports located in the test section model (see Figure 6). Using a light source to illuminate the substance, the flow can then be photographed or filmed with a recording device. The frequency of the light source can be varied in some instances to compensate for faster or slower flow velocities.

4. Water Tunnel Case Studies

Water tunnels have been used as an aerodynamic tool for over a century. They have assisted in the understanding of:

- boundary layers
- flow separation over wings
- vortex shedding
- transition from laminar to turbulent flow
- flow over bluff bodies

An excellent example of water tunnel usage was the vortex flow visualization studies performed by Poisson-Quinton and Werle (Ref.6). The flow situation they were concerned with was the generation of additional lift on slender wings due to vortices forming along the upper surface. These vortices form when the flow separates at the leading edge. Prior tests had shown that the vortex-generated component of the total lift increased as the wing aspect ratio decreased. Poisson-Quinton and Werle used the ONERA water tunnel in France to visualize the different aerodynamic shapes. In virtually each case considered, the water visualization conformed with the actual flow situation. These successes proved that the flow separation off a sharp leading edge was independent of Re .

Another example where a water tunnel assisted in the solving of a problem was at the NAE facility in Canada. Two different designs of a bridge section were tested in the water tunnel: one with edge extensions and one without. The section without edge extensions produced separated flow and a large wake. The section with edge extensions kept the flow attached and produced a smaller wake (thus lower drag). Figure 7 displays these results.

Application to C1 Plenum

The C1 test cell of the Aeropropulsion Systems Test Facility (ASTF) is

currently being modified to support NASP propulsion testing. Based on the projected cost it would require to construct a water tunnel, it has been suggested that the already existing UTSI facility be modified to perform the tests. The UTSI water tunnel would be used to measure the flow quality delivered to the freejet nozzle by the plenum.

1. Description of C1 Plenum

A schematic of the location of the C1 plenum is shown in Fig. 8. There are six concepts being considered for the C1 plenum. Two of the six are illustrated in Figs. 9 and 10. Common to each of these two is the transport of air through a 9 ft. diameter duct into the existing plenum section of the C1. The air is then oriented towards the nozzle inlet by means of a 90° elbow piece. From this point the air is directed to the nozzle with a direct connect device (Fig. 9), or an expanding inner liner (Fig. 10). The other four concepts not discussed here have certain variations in their designs; for a complete description it is recommended Ref. 7 be consulted.

2. Configuration to Water Tunnel

In the configuration process UTSI would be responsible for constructing a plenum, a 90° elbow piece, and a subsonic nozzle. UTSI would also be responsible for setting up and assisting in the running of the tests. The UTSI ductwork has a diameter of 13.5 in.; the existing plenum in the C1 test cell is 22 ft. in diameter. Thus the scaling factor is 0.051, or 5.1%. This small a scale factor raises some concern about the interpretation of data.

Analysis

Problems with Using Water Tunnel on C1

The process air entering the plenum section of the C1 test cell will be at a very low speed; thus, it can be treated as essentially incompressible. This obviates the concern about accounting for compressibility effects in the water

model. However, the C1 inlet/plenum concepts present some unique problems for a water tunnel analysis:

- While water tunnel visualization has been used to study a variety of phenomena, one characteristic has been common to virtually every successful effort. This characteristic is prior knowledge of the flow structure and any anomalies associated with it. The distinctions between air and water outlined earlier force the experimenter to understand the actual flow before it is modelled in a different medium. Otherwise, erroneous results could be obtained without knowing they are erroneous. Terri Brock, a Sverdrup test engineer working on the DDO3 project, has pointed out that there is no knowledge of a problem with the plenum flows. Without knowing what to look for, this causes the evaluation of the various plenum concepts to be very difficult.
- The AEDC report specifies certain quantitative criteria for the inlet and exit planes of the nozzle (i.e., pressure and temperature ranges, flow distortion levels, and expected boundary layer thicknesses). Using a water tunnel model to measure and account for these specifications could lead to serious mistakes in the obtained results. The inability to acquire accurate quantitative data from a water tunnel again lies in the fact that water is being used to simulate air. The strength of the water tunnel is in its ability to make a flow visible, thus giving a person an intuitive grasp of the situation. No extensive research exists documenting successful quantitative analysis with a water tunnel.
- In the upstream piping of the ASTF complex, there is a series of four 90° bends which the air must travel through before reaching the plenum area. The current proposed mechanism for testing the plenum concepts does not incorporate these bends. Dr. Frank Collins, an aerospace engineering professor at UTSI, has done extensive work with water tunnels. In an experiment he performed, inconsistencies were discovered in the results between the water

model and the actual wind tunnel. The problem was eventually test section. In the case of ASTF, the upstream bends could pose problems with separated flow. A study on flow through right angle bends (Ref. 8) examined the nature of flow separation in them. The study demonstrated that the region of separated tended to increase in size as the approach Mach number increased (Fig. 11). Also, the flow did not appear to completely reattach even at significant downstream distances. A water tunnel analysis not accounting for the upstream right angle bends of ASTF could give misleading results.

Recommendations and Conclusions

I think a water tunnel analysis of the C1 plenum concepts poses problems which will make the evaluation process highly prone to error. Water tunnels are undoubtedly excellent tools for flow visualization, and can give the engineer or scientist an added sense of understanding. However, their ability to yield accurate quantitative information is skeptical. Coupled with this are the unaccounted for upstream bends and the lack of prior knowledge of any flow problems. Another test technique, such as CFD or subscale analysis, will likely prove more beneficial. However, water tunnel testing could be valuable as a supportive technique in later test stages.

REFERENCES

1. Dobrodzicki, G.A., "Flow Visualization in the National Aeronautical Establishment's Water Tunnel", Aeronautical Report LR-557, National Research Council of Canada, February 1972.
2. Matthews, C.W., "The Design, Operation, and Uses of the Water Channel as an Instrument for the Investigation of Compressible Flow Phenomena", NACA Technical Note 2008, Washington, January 1950.
3. McDaniels, D.M., "Background and Theoretical Considerations for Utilizing Water Tunnel Flow Visualization", Master's Thesis, UTSI, August 1986.
4. Lehman, A.F., and Peirce, T.E., "The Garfield Thomas Water Tunnel-A Tool for Hydrodynamic Research", 17th American Rocket Society Meeting, Los Angeles, California, November 13-18, 1962.
5. Collins, F.G., "Description of the UTSI Water Tunnel", UTSI Report 82-1, UTSI, January 1982.
6. Poisson-Quinton, P., and Werle, H., "Water Tunnel Visualization of Vortex Flow", Astronautics and Aeronautics (Vol.5), pp.64-66, June 1967.
7. Dillehay, G.W., "Test Cell C-1 Modifications to Support NASP Testing", AEDC-TMR-89-E29, December 1989.
8. Ranov, T., and Graves, K.W., "Separation of Channel Flow Induced by a Right Angle Bend", ASME Symposium on Fully Separated Flows, ASME Fluids Engineering Division Conference, Philadelphia, Pennsylvania, May 18-20, 1964.

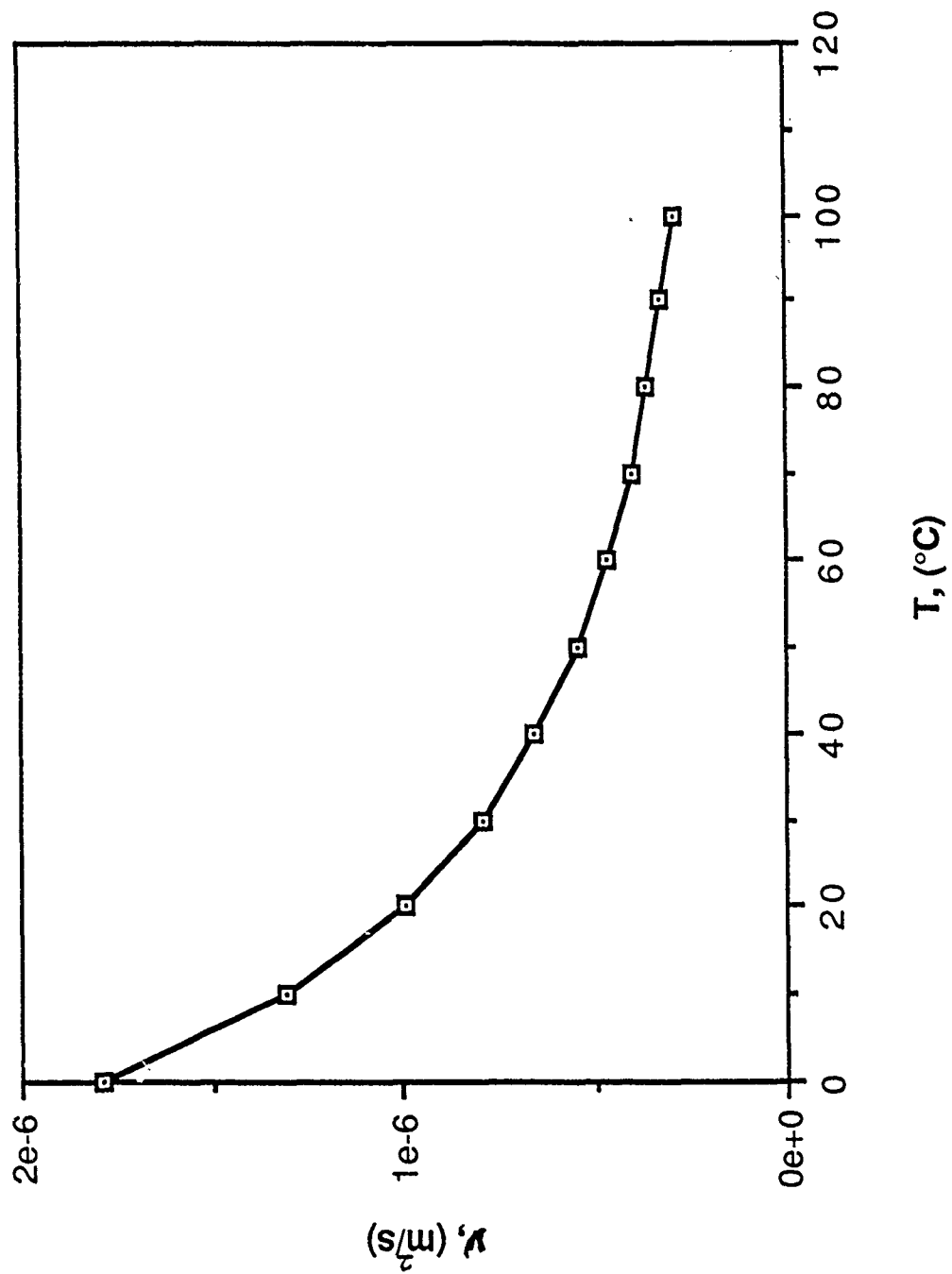
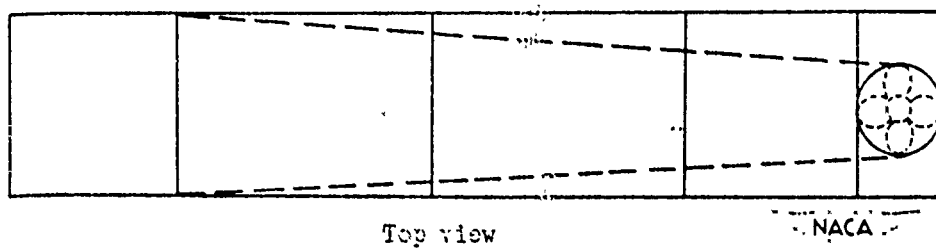
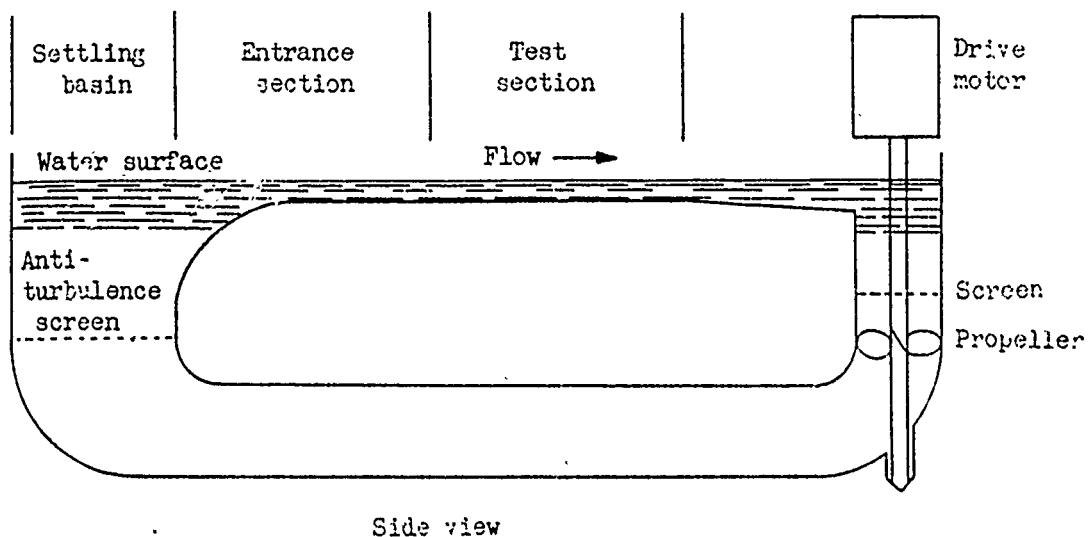


Fig. 1 Water Viscosity Variance with Temperature



(a) Vertical return channel.

Fig. 2 Vertical Water Channel (Ref. 2)

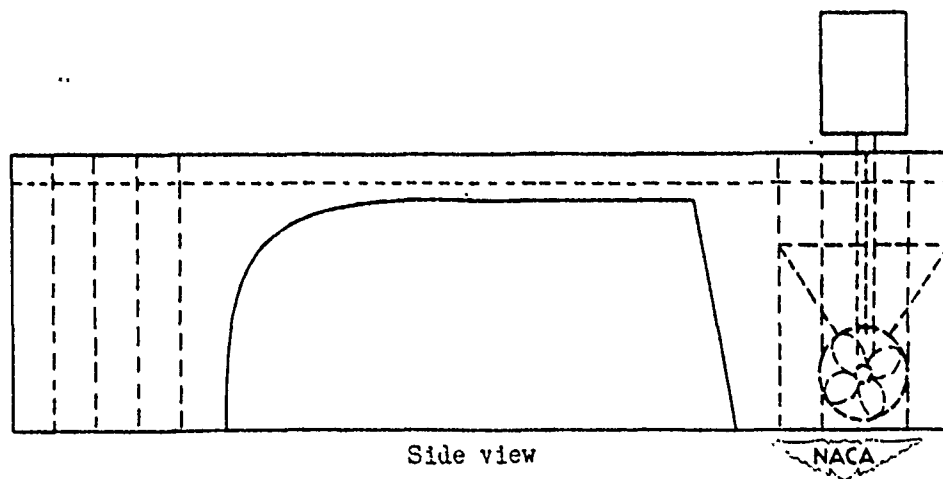
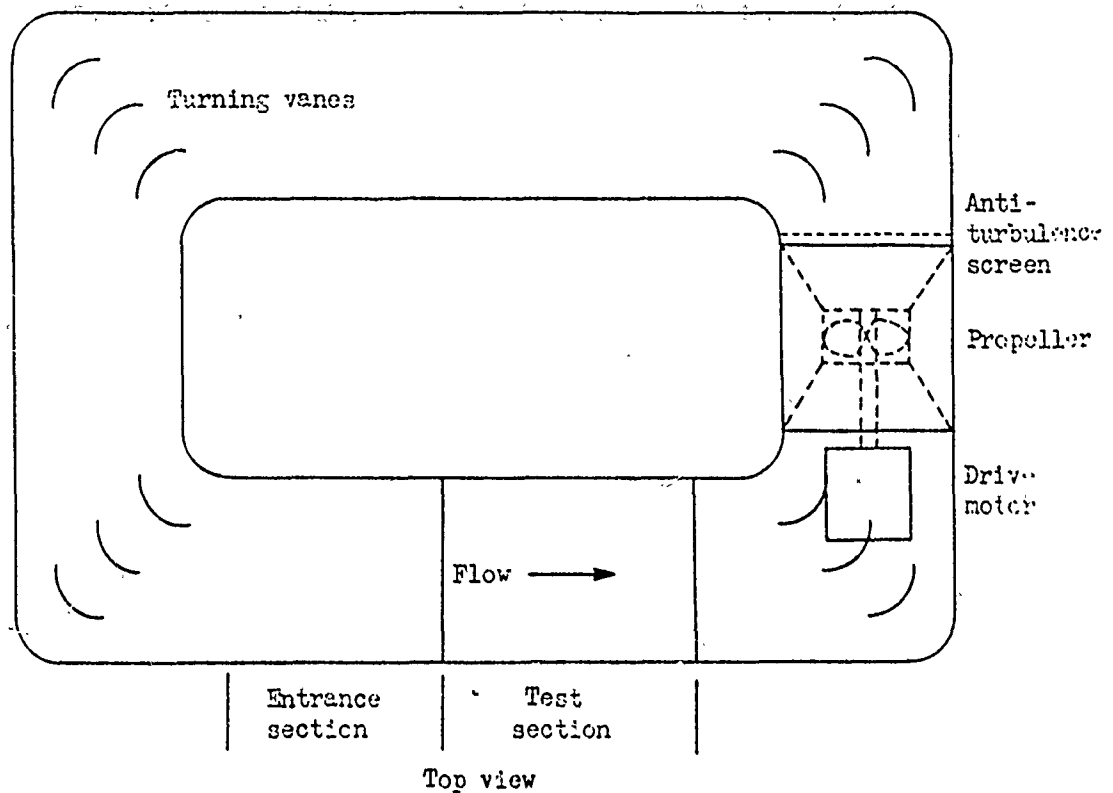


Fig. 3 Horizontal Water Channel (Ref. 2)

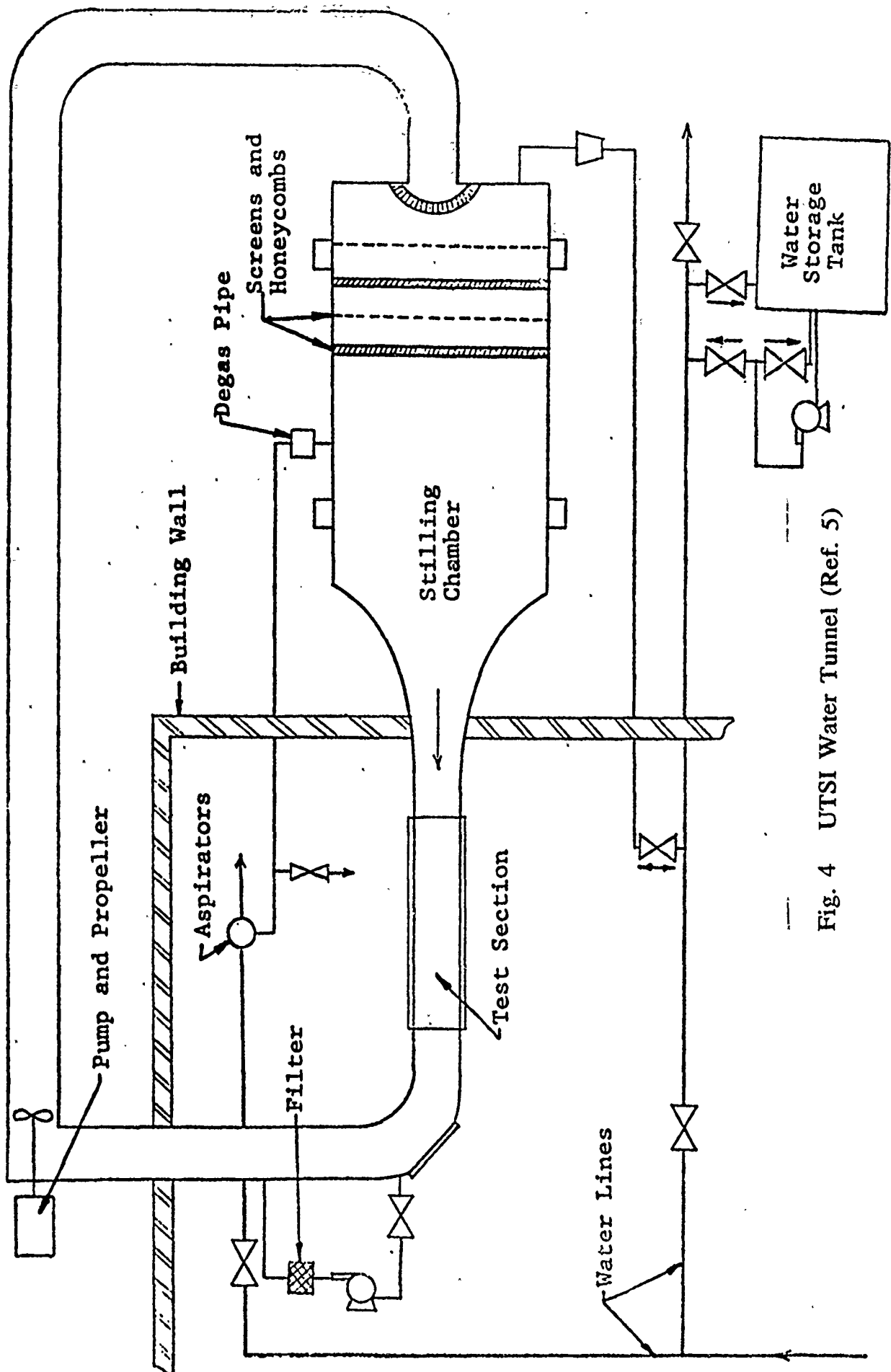


Fig. 4 UTSI Water Tunnel (Ref. 5)

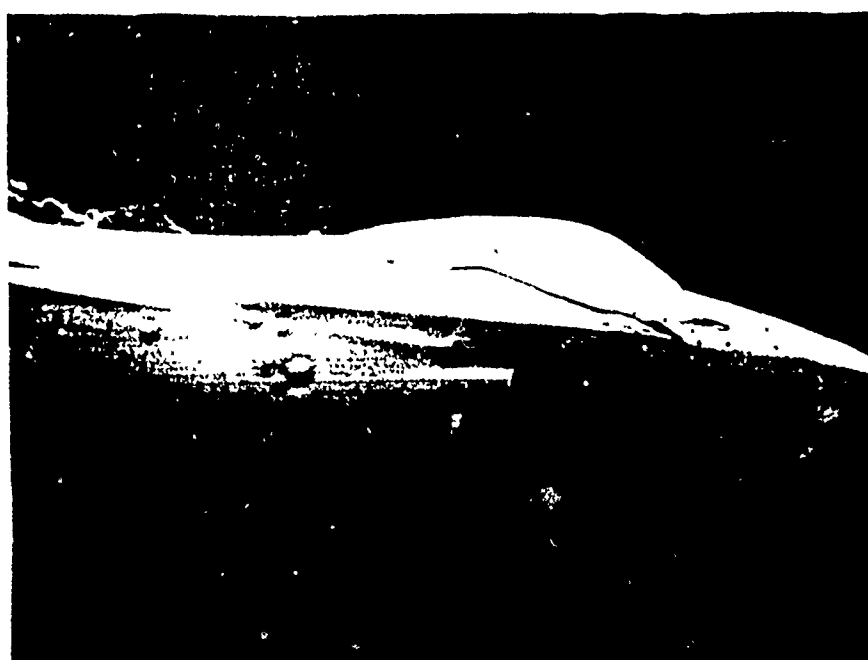
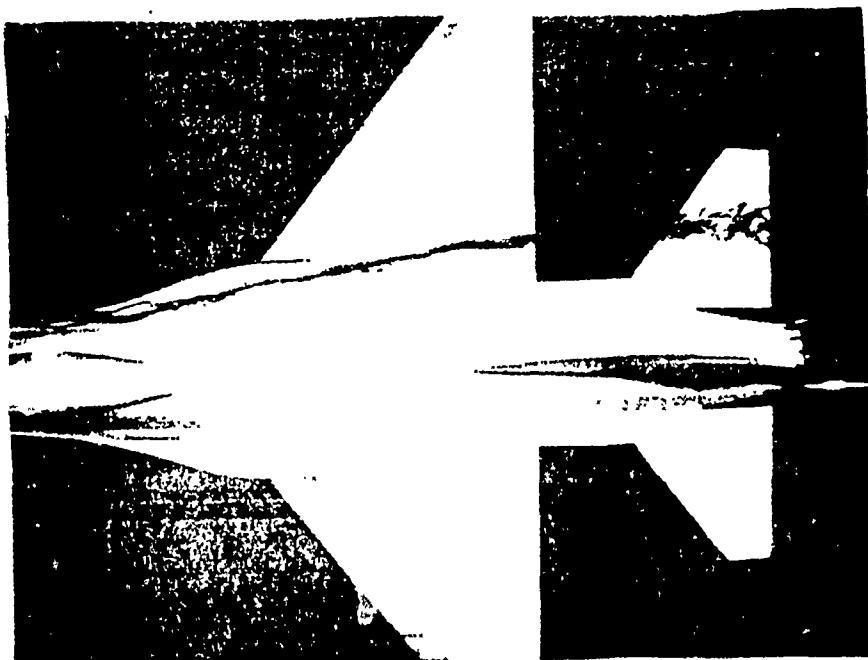
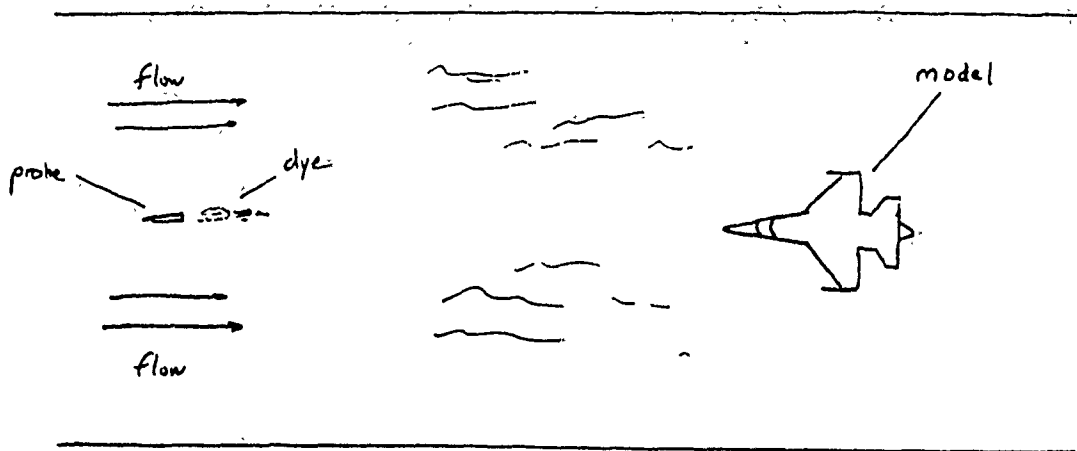
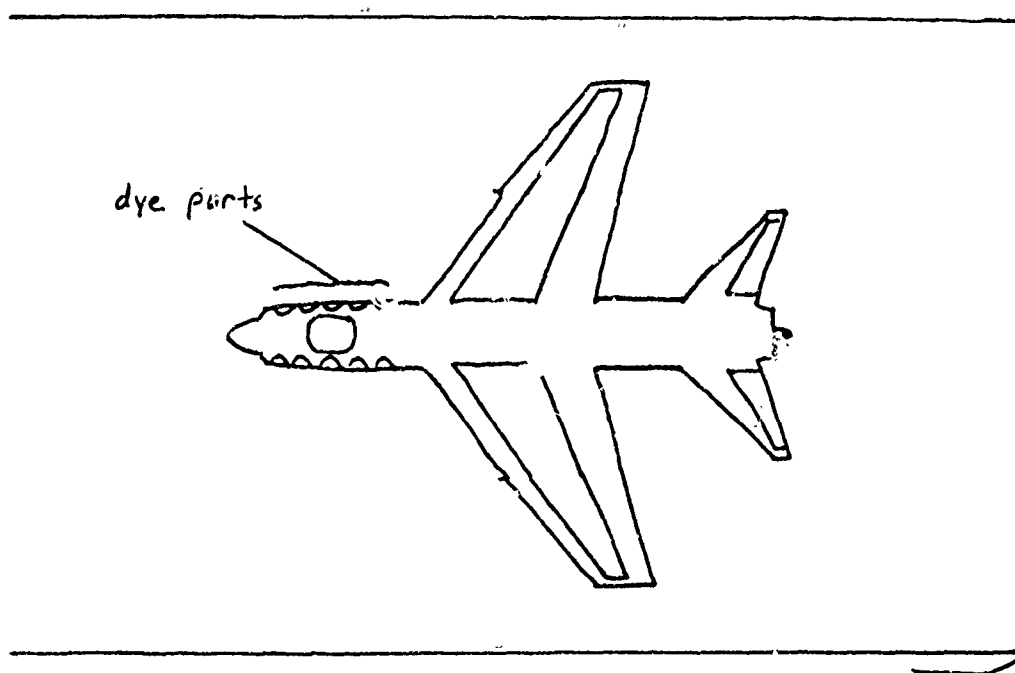


Fig. 5 Strake Vortex Flow on an F-15 (Ref. 5)

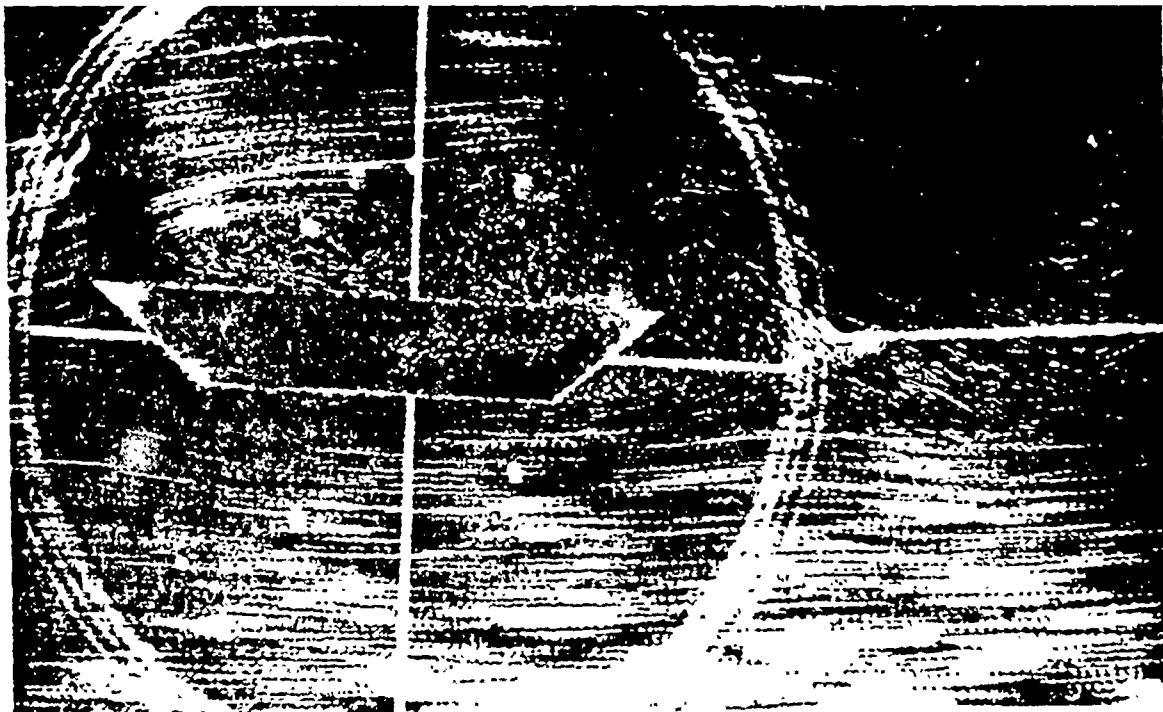


a) Upstream probe

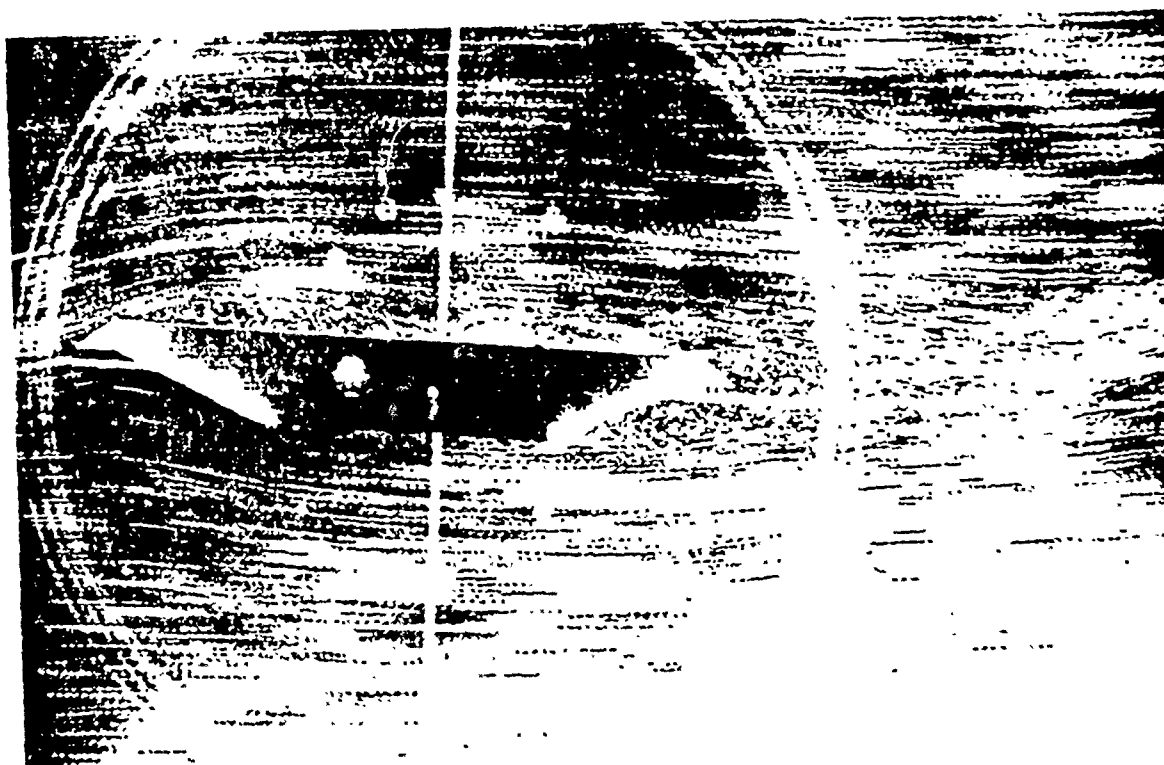


b) Injection through model ports

Fig. 6 Methods of Dye Injection



ORIGINAL SECTION



IMPROVED SECTION (EDGE EXTENSION)

Fig. 7 Water Tunnel Flow Over a Model Bridge Section

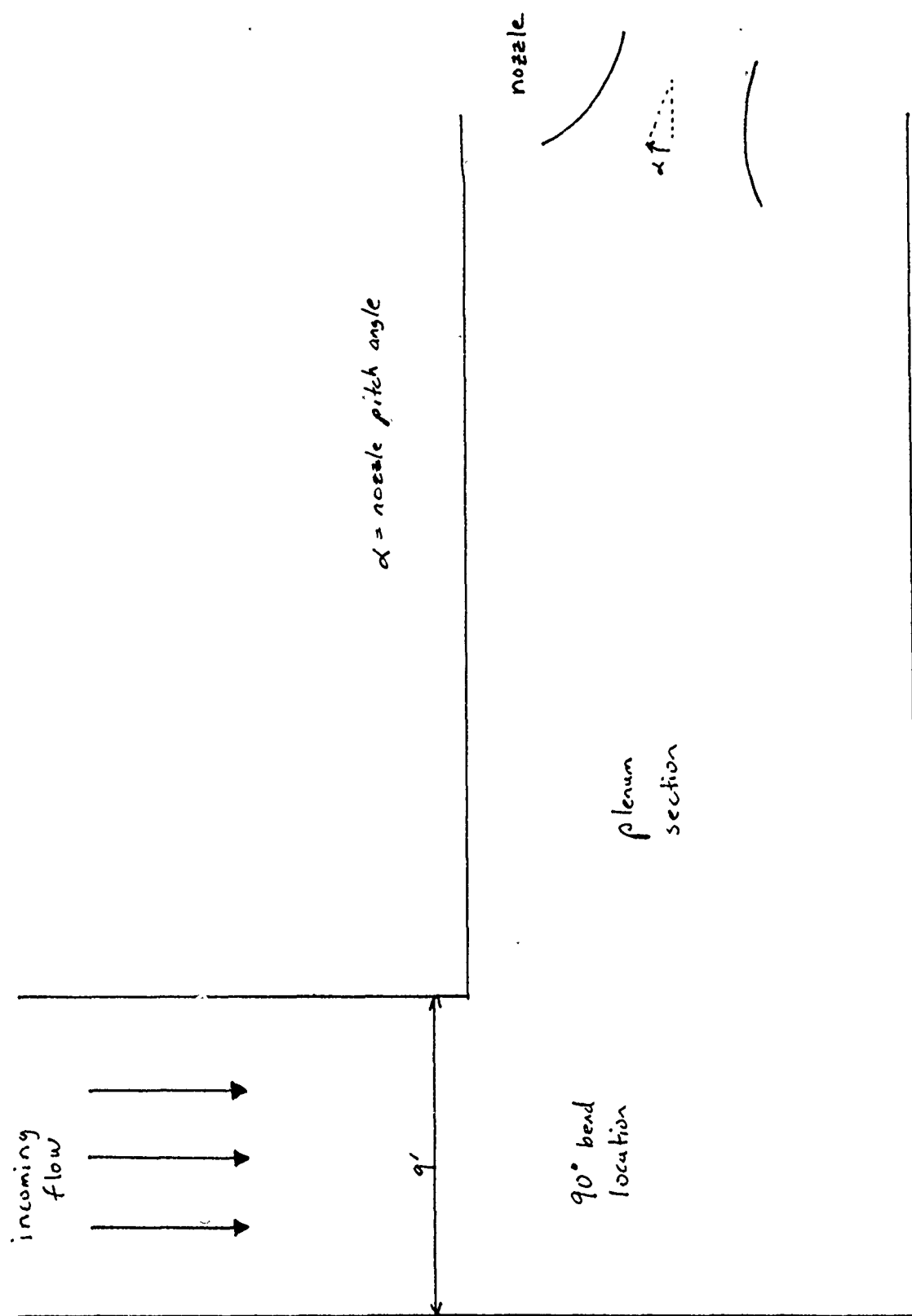


Fig. 8 C1 Plenum

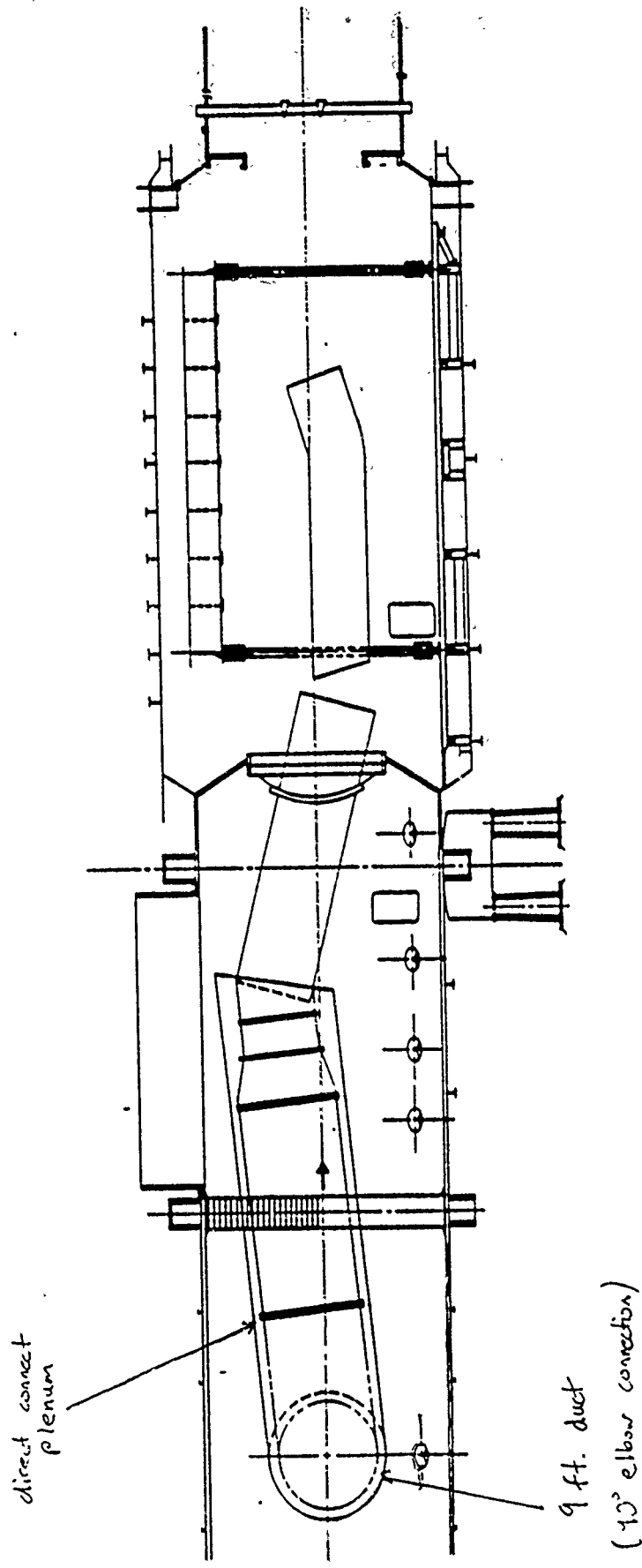


Fig. 9 Direct Connect Concept (Ref. 7)

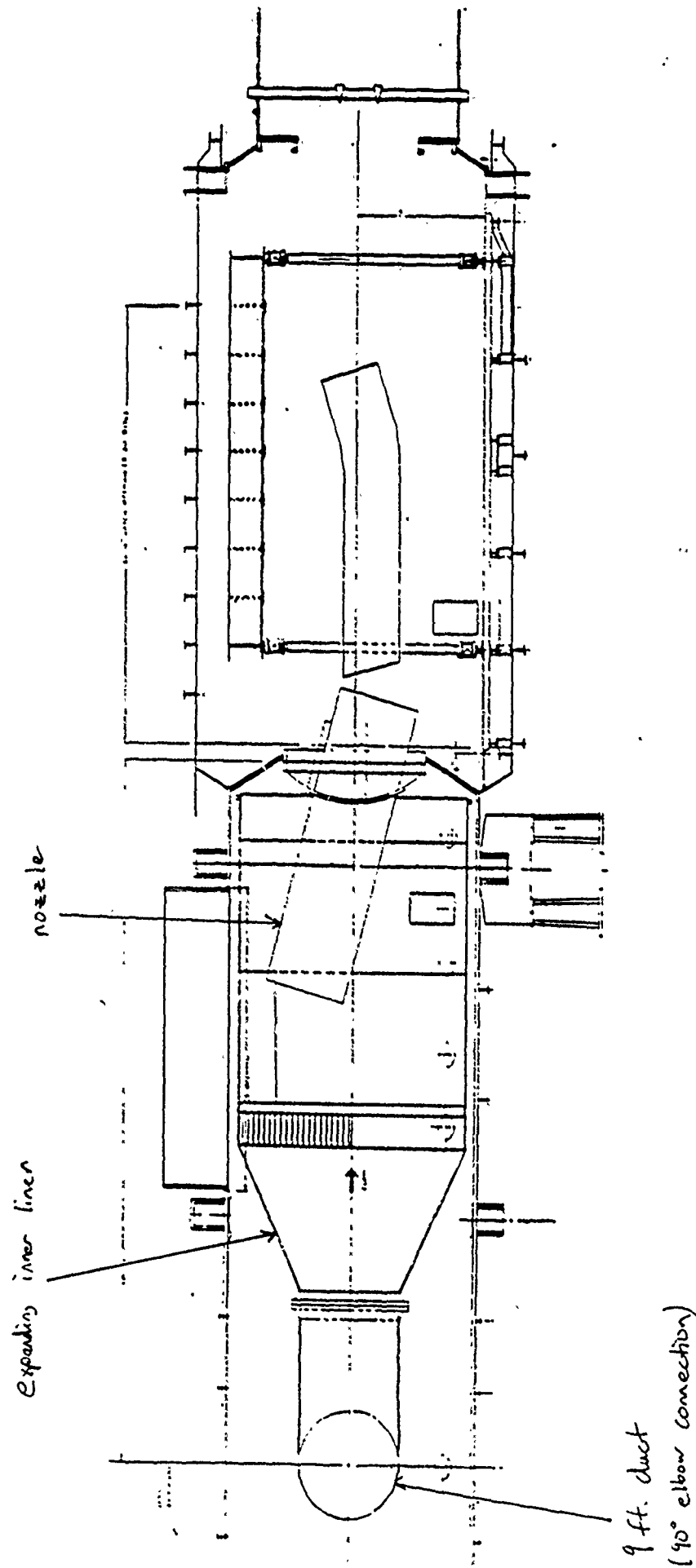


Fig. 10 Expanding Inner Liner Concept (Ref. 7)

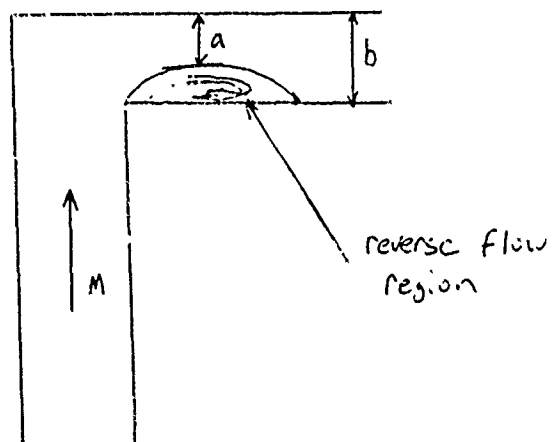
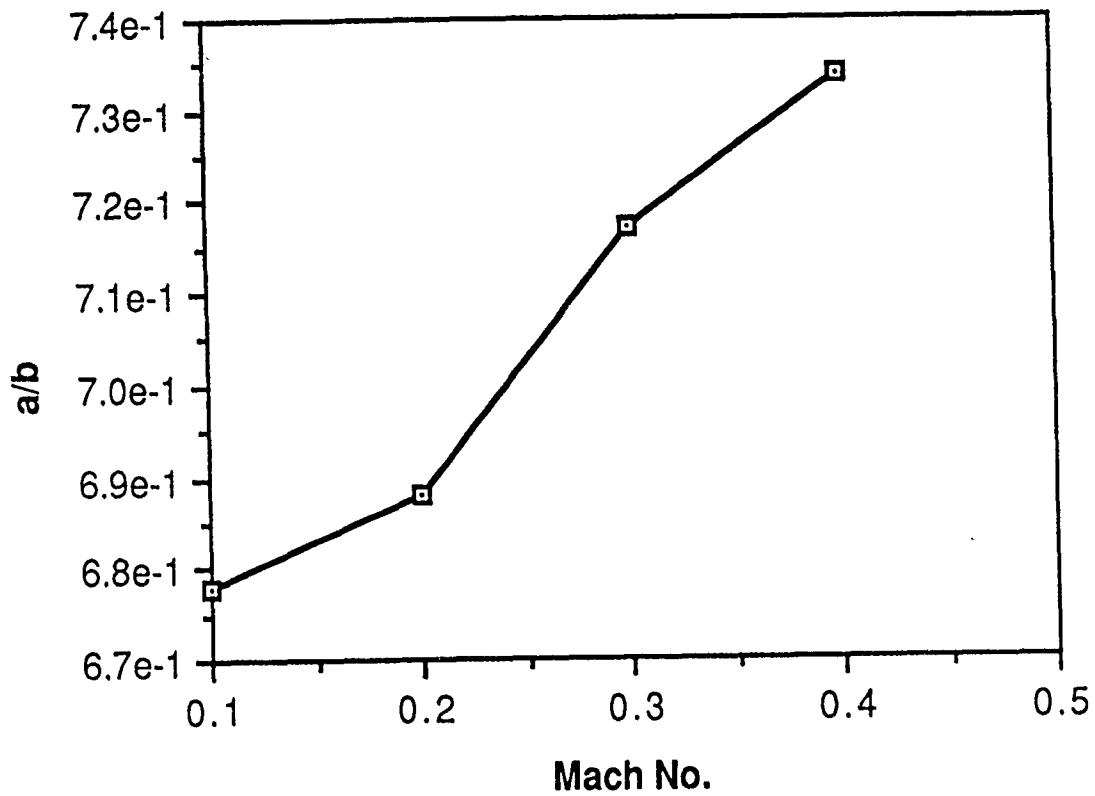


Fig. 11 Reverse Flow Region in a Right Angle Bend
(Data from Reference 8)

CALIBRATION TECHNIQUE FOR COCODEC

Emily G. Joy
Graduate Research Associate
Arnold Engineering Development Center

ABSTRACT

This paper presents a calibration procedure for a compressor flow code called COCODEC. The technique involves adjusting relative inlet dynamic head loss coefficient and relative exit angle deviation within the code to match calculated values of overall pressure ratio and efficiency with experimental results for a single rotor and a three stage transonic fan. A method of correlating the data from the calibration of each machine is discussed. Performance correlations for each machine are formulated for intended implementation into a three-dimensional flow code.

INTRODUCTION

The Propulsion Computational Technology Section at Arnold Engineering Development Center (AEDC) has determined the necessity of achieving accurate stage-by-stage performance maps for a transonic compressor using a steady state compressor flow code. The code available for this purpose is the Combined Compressor Design Code (COCODEC), an axisymmetric, steady state, flow code with the capability of modeling the flow over a compressor blade row. An input file provides the blade geometry. This geometry defines the flow path upstream and downstream of the compressor as well as the overall characteristics of each blade row.

COCODEC contains loss and deviation correlations developed in the 1950's and 1960's associated with the NASA 65 series of airfoils and double-circular-arc, two-dimensional cascades. The data used in the correlations were obtained in the subsonic

operating regime. However, the blading used in current turbomachinery applications is designed for transonic flow. Therefore, although the proper geometry can be included through the input file, the old correlations do an inadequate job of predicting the performance of compressors designed with transonic blading. The objective of this project was to calibrate the existing code to yield more accurate predictions of the overall compressor pressure ratio and adiabatic efficiency. The approach was to systematically adjust the performance of the rotor blades at each stage. These adjustments were then mapped and collapsed to find more appropriate correlations. This calibration technique was employed for two machines for which experimental data was available for comparison. The first was a single rotor, called Rotor 2D (Ref. 1), and the other was a three stage transonic fan with three sets of rotors and stators and an inlet guide vane (Ref. 2).

CORRECTION METHODOLOGY

There are several variables included in COCODEC that can be used to adjust the flow across the blade rows. The two of interest for this calibration were ADD LOSS and ADD DEVIATION. ADD LOSS adds or subtracts a percentage of the inlet dynamic head total-pressure-loss coefficient across a blade row. ADD DEVIATION causes the angle of the flow to be smaller or larger than the expected deviation from the metal angle at which the flow leaves the blade. The ADD LOSS percentage and the ADD DEVIATION angle can be input as arbitrary values at each stage. The output yields the corresponding overall characteristics for the machine, including total pressure ratio and rotor adiabatic efficiency. By changing ADD LOSS and ADD DEVIATION a small increment at a time, the pressure ratio and efficiency can be matched to the experimental data.

Two different compression systems were adjusted using this method, a single stage rotor (Rotor 2D) and a three stage transonic fan. The adjustments and corresponding

correlations are discussed in the next sections.

Rotor 2D

The performance map for this rotor is given in the Appendix (Ref. 1). For each combination of corrected weight flow and corrected speed, a total pressure ratio and rotor adiabatic efficiency has been mapped from experimental results. The point at the extreme left of each speed line is the stall point. The extreme right of these lines approaches choked flow, and the characteristics become nearly vertical at high speeds.

COCODEC was run for numerous trials until the appropriate values of ADD LOSS and ADD DEVIATION could be determined to match the calculated and the experimental results of pressure ratio and efficiency, simultaneously, at a designated weight flow and corrected speed. There was a unique combination of ADD LOSS and ADD DEVIATION that generated the calculated values that were closest to the experimental results. Therefore, this procedure has the potential for automation with an optimization scheme. A sample comparison of how closely the experimental data and calibrated calculations matched is shown in Figures 1 and 2 below. Figure 1 maps the total pressure ratio vs. the corrected weight flow in lb/sec for the seventy percent corrected speed. Figure 2 plots the rotor adiabatic efficiency vs. the corrected weight flow for the same speed line. The values produced by COCODEC, using an appropriate ADD LOSS and ADD DEVIATION to calibrate it, came very close to the experimental results with occasional exceptions. The pressure ratio and efficiency proved to be accurate to at least two decimal places when the optimum ADD LOSS and ADD DEVIATION were found.

For each speed line, a set of ADD LOSS and ADD DEVIATION combinations that yielded the closest pressure ratios and efficiencies was recorded and plotted against the corresponding weight flows to examine the trends. Each of the adjustments indicated

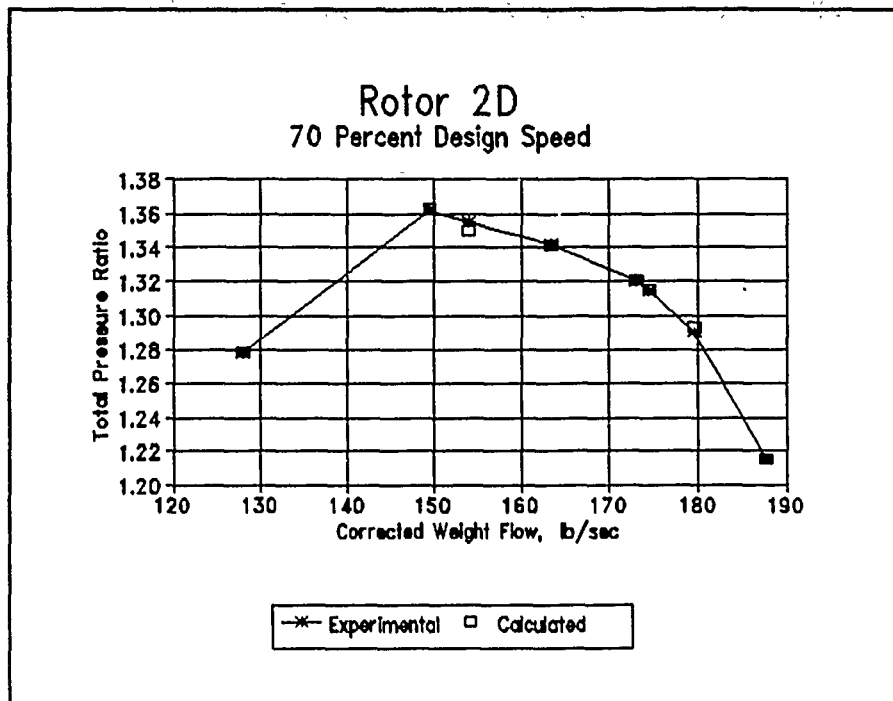


Figure 1 Comparison of Experimental & Calculated Pressure Ratio for Rotor 2D

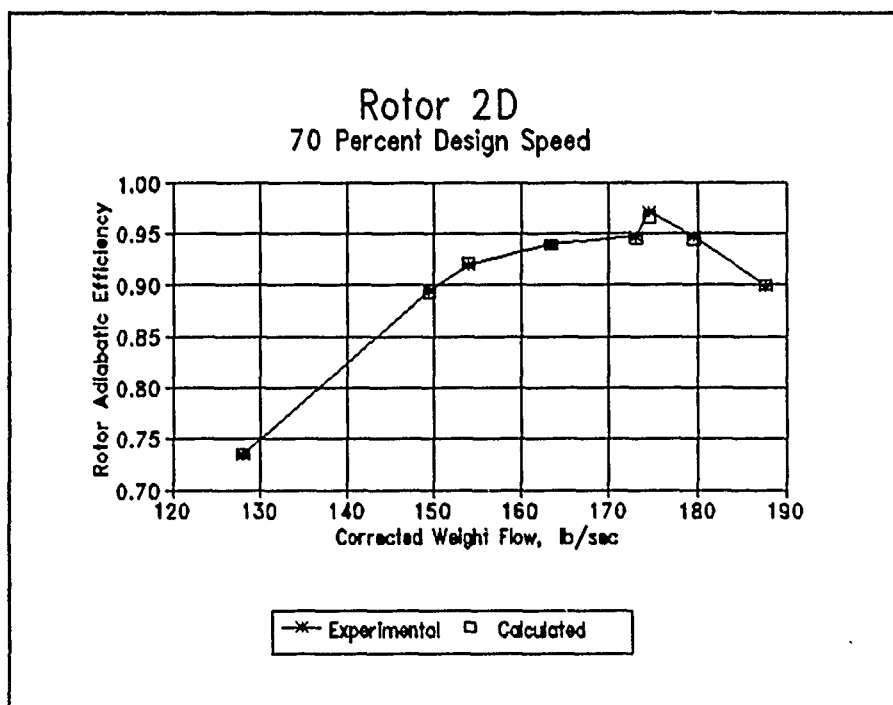


Figure 2 Comparison of Experimental & Calculated Adiabatic Efficiency for Rotor 2D

similar patterns from speed to speed when plotted. As an example, the resulting ADD LOSS values are shown in Figure 3. The values on the extreme left of each speed correspond to the stall point. Each of these fell outside of the parabolic shape associated with that speed and required the greatest adjustment within that flow range. This could result from the uncertainty of the experimental data.

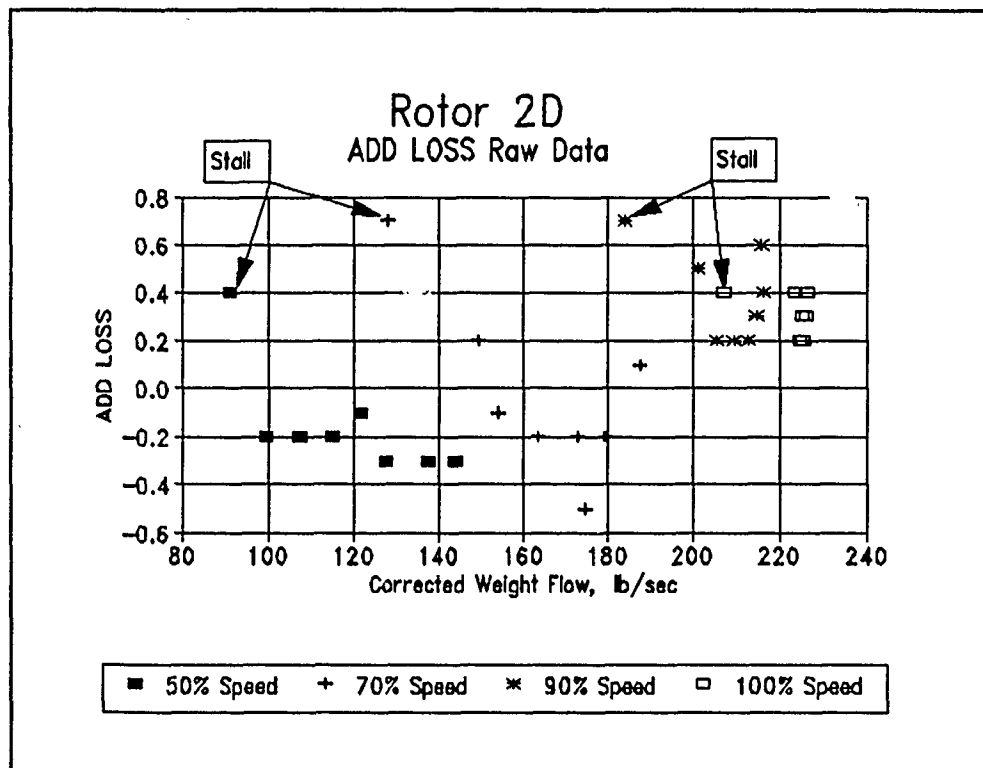


Figure 3 ADD LOSS Raw Data at Various Speeds for Rotor 2D

The next task involved correlating this data into a form that could be easily implemented into the code. A collapsing process developed by William B. Roberts (Ref. 3) was utilized to accomplish this task to avoid a complicated interpolation scheme. The first step entailed normalizing the weight flow to move the flow range for each speed into a single range on the x-axis. This normalized flow parameter, w_{norm} , is given by Equation 1.

$$w_{norm} = \frac{w - w^*}{w_{max} - w_{min}} \quad (1)$$

Here, w is the corrected weight flow, w_{max} is the maximum weight flow in the range, w_{min} is the minimum weight flow in the range, and w^* is the average of the maximum and minimum given by Equation 2.

$$w^* = \frac{w_{max} + w_{min}}{2} \quad (2)$$

The middle of the flow range has been chosen since a compressor frequently operates near this point. Normalizing the weight flow to a non-dimensional value in Equation 1, the curves of Figure 3 overlap as shown in Figure 4. This normalization was an improvement,

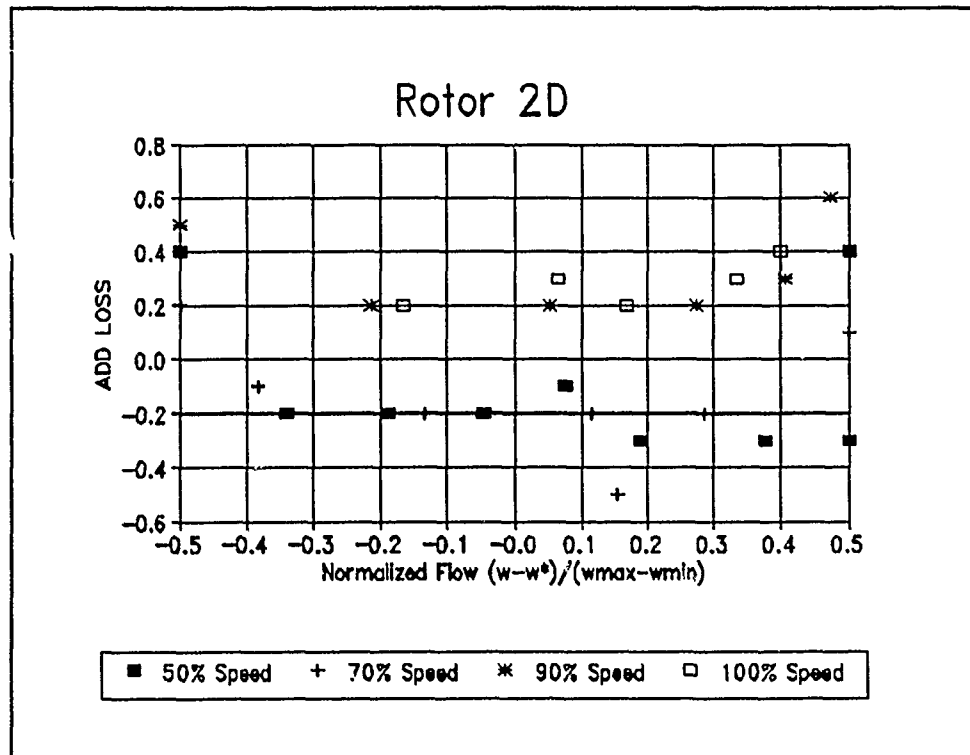


Figure 4 ADD LOSS Data with a Normalized Flow Parameter for Rotor 2D

but the data spanned a fairly wide band in the y-direction. It was difficult to get a single curve fit that sufficiently modelled the data. Therefore, another step was necessary. To bring the curves closer together in the y-direction, each curve was zeroed with respect to its mean value. In other words, the arithmetic mean of each range of data was subtracted from each data point within that range. This placed the mean of each range at zero on the y-axis and referenced all other points to it. The results of this collapsing procedure are shown in Figure 5 below. After this step, the data was close enough that a reasonable curve

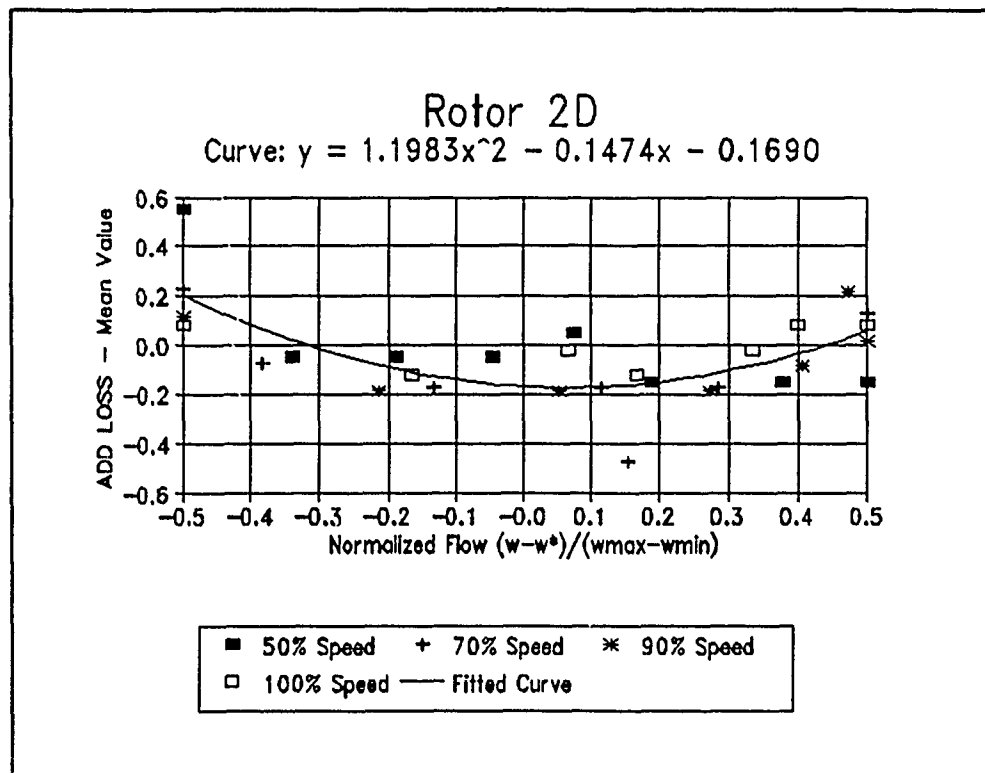


Figure 5 Collapsed ADD LOSS Data & Curve Fit for Rotor 2D

fit was possible. The extreme values still showed some scatter, but these points represented the areas where measurement became more uncertain at stall and choked flow. A quadratic curve, given in Equation 3, has been fit through this data set using the method of least squares to yield a useful correlation to implement into the code.

$$ADD\ LOSS - Mean = 1.1983 \times w_{norm}^2 - 0.1474 \times w_{norm} - 0.1690 \quad (3)$$

The same technique was used on the ADD DEVIATION data to collapse it. Both the normalization of the x-axis and the zeroing of the mean ADD DEVIATION value were necessary to generate a narrow band of data. Figure 6 demonstrates that this procedure

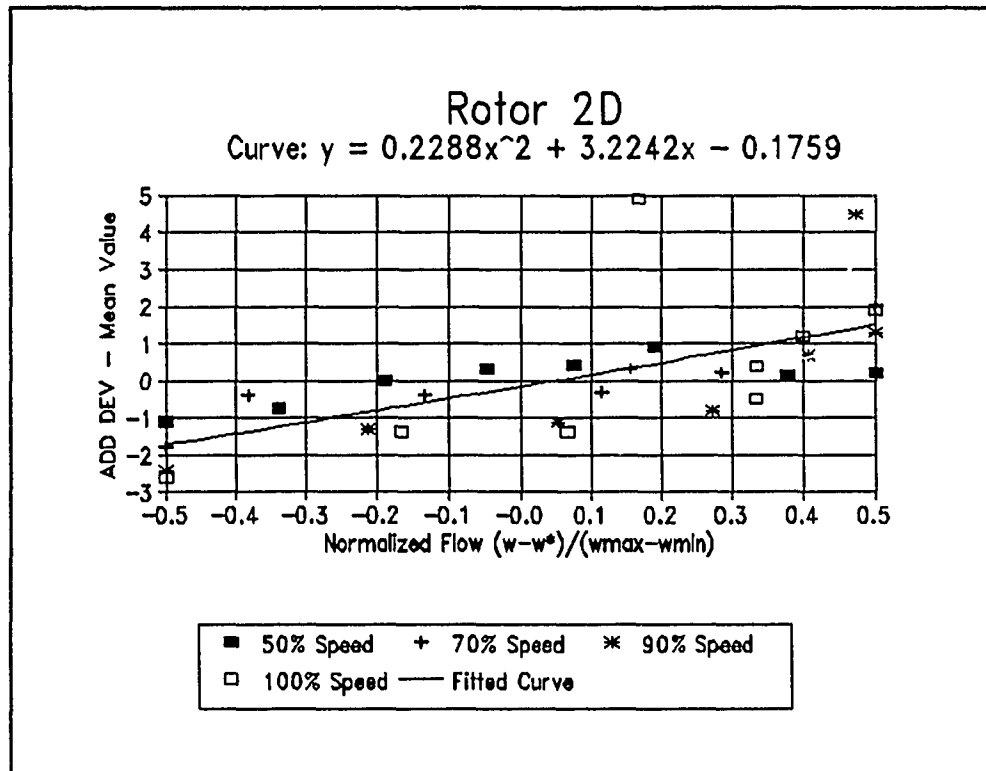


Figure 6 Collapsed ADD DEVIATION Data & Curve Fit for Rotor 2D

was effective in collapsing the data so that it could be fit to a quadratic relation given in Equation 4.

$$ADD\ DEV - Mean = 0.2288 \times w_{norm}^2 + 3.2242 \times w_{norm} - 0.1759 \quad (4)$$

Three Stage Transonic Fan

The performance map for this machine is also shown in the Appendix (Ref. 2). This data represents the overall characteristics for the three stages together. The stall points are not included for this machine, but each curve does become vertical as the choked end is approached.

The procedure to obtain the correct pressure ratio and efficiency was fundamentally the same. However, since this machine consisted of three stages as opposed to one, there was an ADD LOSS and ADD DEVIATION associated with each stage. The experimental characteristics were only available for the overall performance as opposed to stage-by-stage performance. In order to match the data given, the adjustments were set at the same constant value for each rotor with the stators unadjusted in each trial. This resulted in a unique solution for each speed and weight flow which may not have been possible if each stage was allowed a different set of ADD LOSS and ADD DEVIATION values. Again, this procedure may lend itself to automation. A sample comparison of how closely the experimental data and calibrated calculations matched for this machine is shown in Figures 7 and 8 below. Figure 7 plots the total pressure ratio vs. the corrected weight flow for the ninety percent corrected speed. Figure 8 is a plot of the corresponding adiabatic efficiency vs. the corrected weight flow at this speed. Again, the corrected calculations matched the experimental results with suitable accuracy.

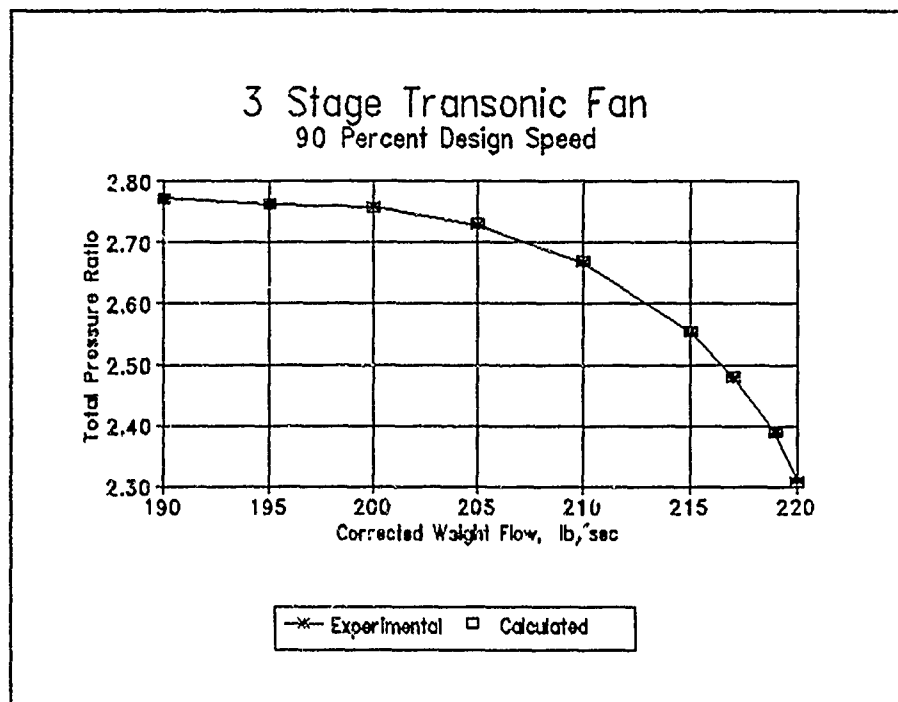


Figure 7 Comparison of Experimental & Calculated Pressure Ratio for a 3 Stage Fan

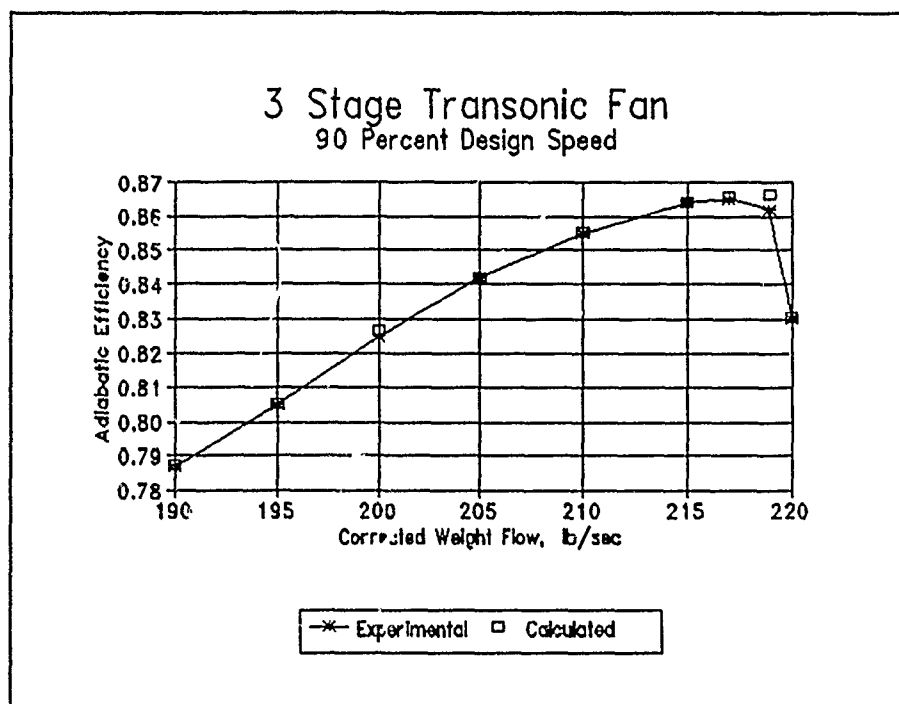


Figure 8 Comparison of Experimental & Calculated Adiabatic Efficiency for a 3 Stage Fan

The calibration variables for this case were also plotted against the appropriate weight flow to view the trends. The ADD LOSS data is shown in Figure 9. The trend in this figure was evident, but it was very different from that of the single rotor. For Rotor 2D, the losses could only be attributed to the rotor, whereas there were rotors and stators associated with each stage in the fan. In addition, there was an inlet guide vane on the fan which was set at a specific value for each speed.

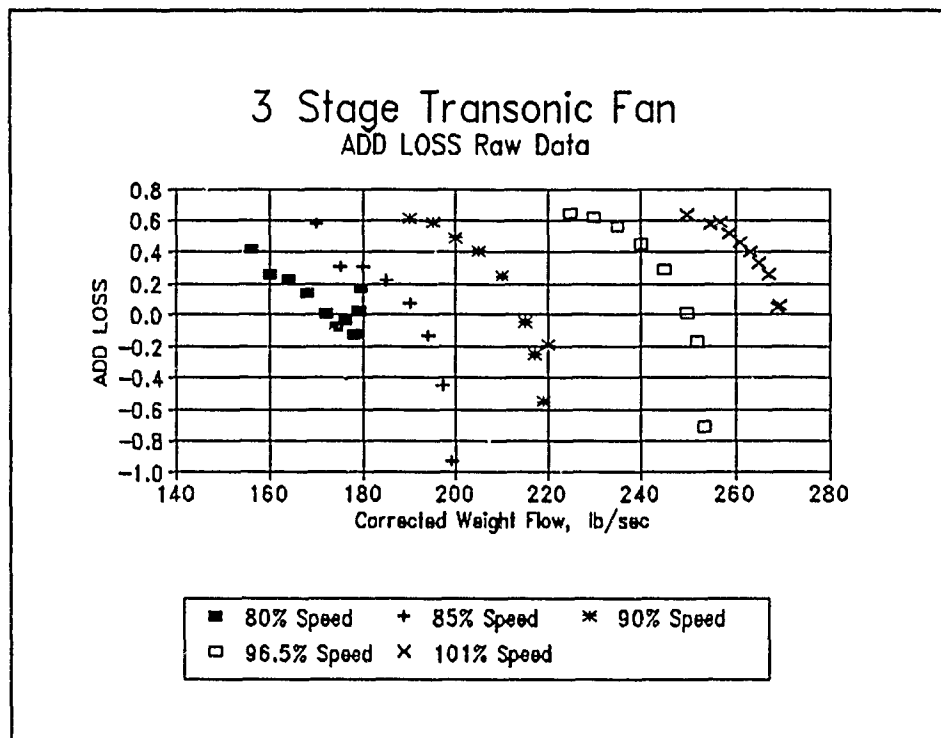


Figure 9 ADD LOSS Raw Data at Various Speeds for a 3 Stage Fan

The same collapsing method used for Rotor 2D was employed with the three stage transonic fan. The results of normalizing the data along the x-axis using Equations 1 and 2 above are displayed in Figure 10. Since there are no stall points to consider in this case, the widest scatter appeared on the choked end only. By zeroing the mean of the range of each speed line, the data was collapsed to the form of Figure 11. A quadratic curve fit

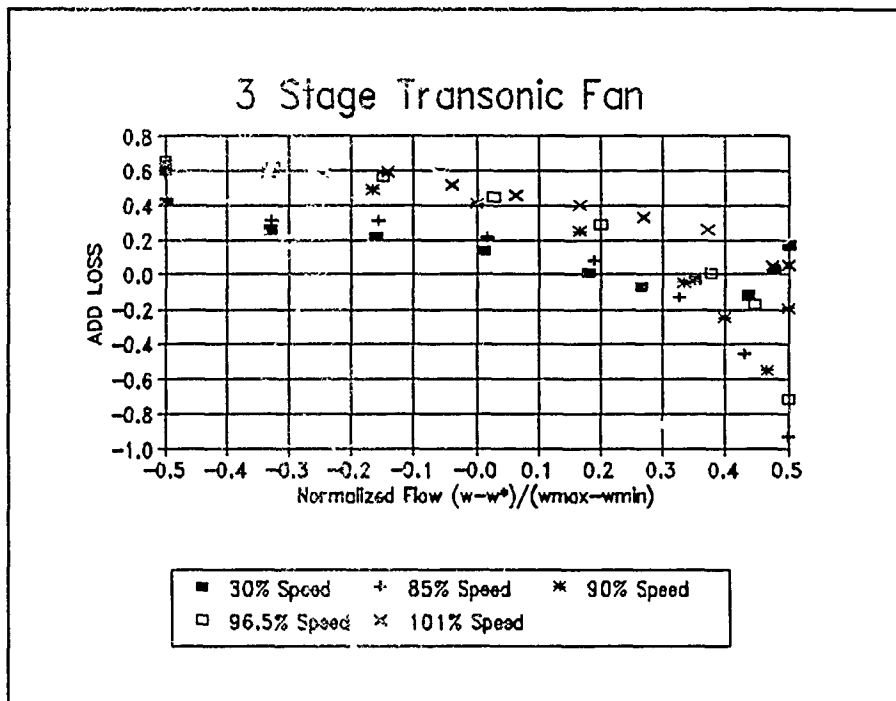


Figure 10 ADD LOSS Data with a Normalized Flow Parameter for a 3 Stage Fan

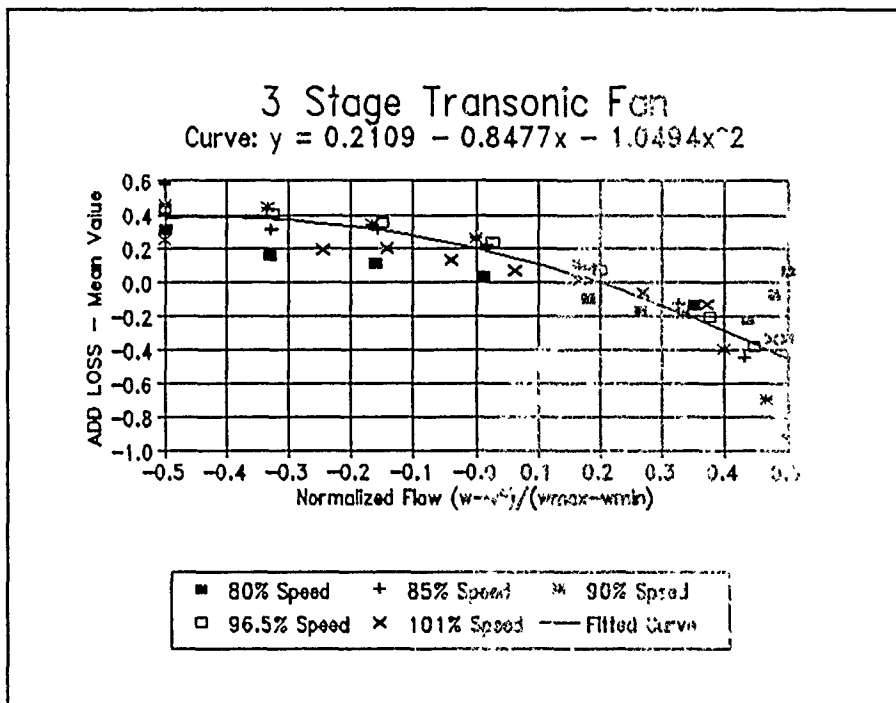


Figure 11 Collapsed ADD LOSS Data & Curve Fit for a 3 Stage Fan

using the method of least squares was generated. The ADD LOSS correlation is given in Equation 5.

$$ADD\ LOSS - Mean = -1.0494 \times w_{norm}^2 - 0.8477 \times w_{norm} + 0.2109 \quad (5)$$

For ADD DEVIATION, the trends in the raw data for this adjustment also differed from that for Rotor 2D. However, the data collapsed nicely to a narrow band as shown in Figure 12. Collapsing was necessary on both axes to achieve a good curve fit. The

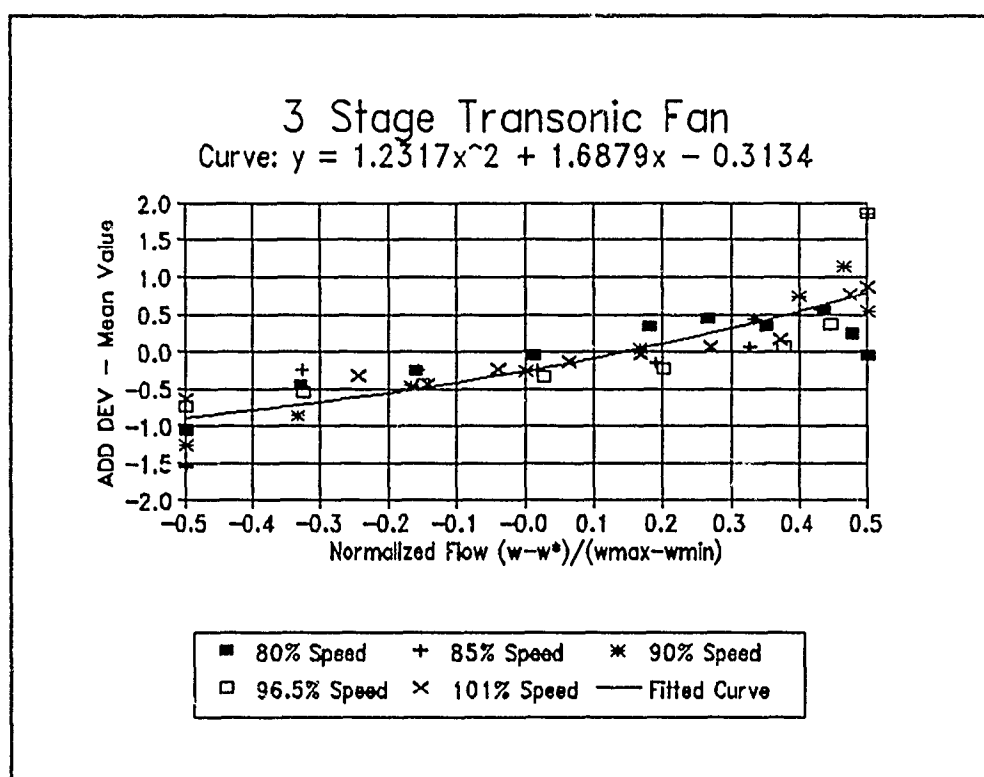


Figure 12 Collapsed ADD DEVIATION Data & Curve Fit for a 3 Stage Fan

correlation that fit this collapsed data with the least error is given in Equation 6.

$$ADD\ DEV - Mean = 1.2317 \times w_{norm}^2 + 1.6879 \times w_{norm} - 0.3134 \quad (6)$$

CONCLUSIONS

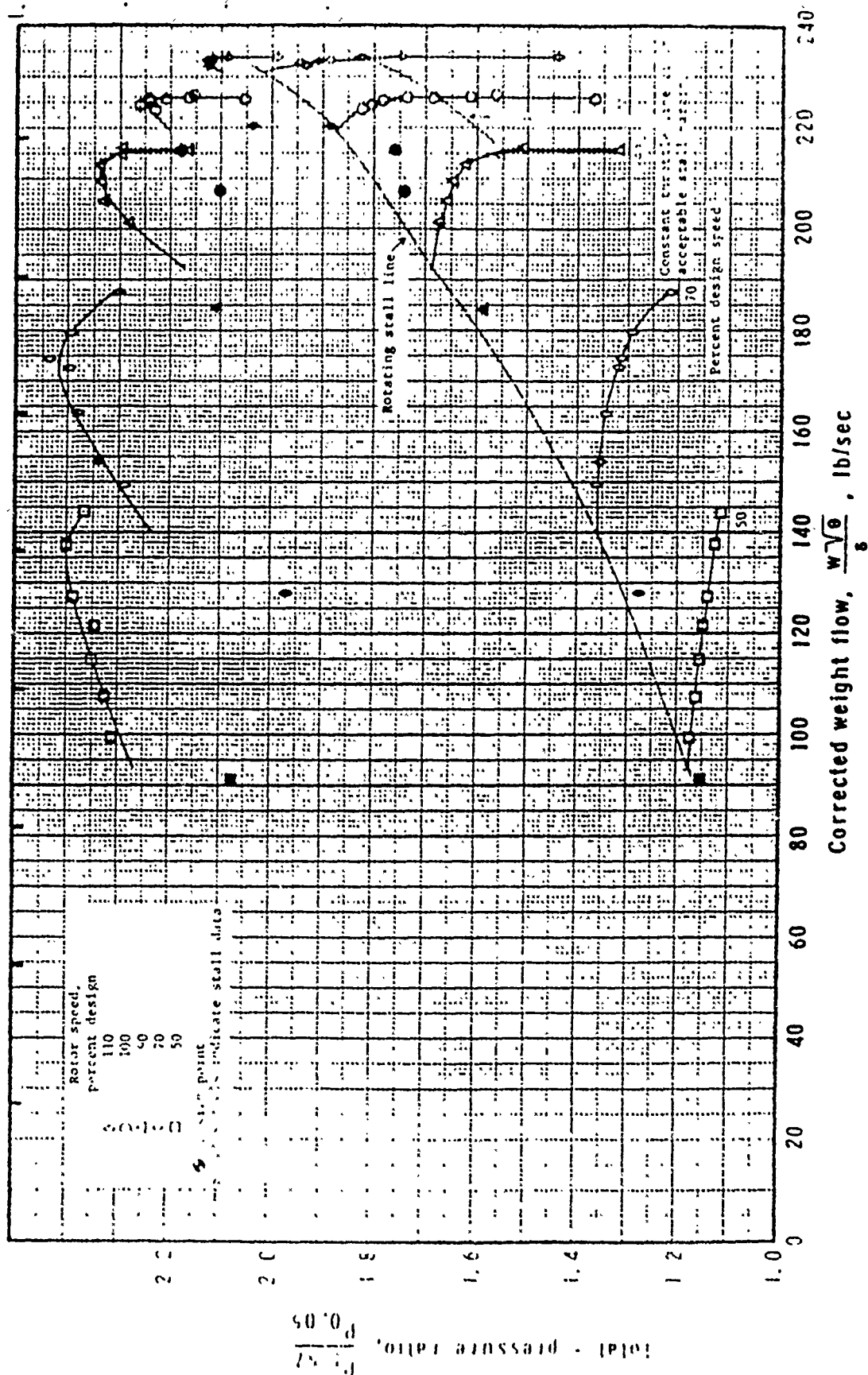
Performance correlations for COCODEC have been formulated for a single rotor (Rotor 2D) and for a three stage transonic fan. Each machine required an independent correlation. However, the single rotor could be expected to have different losses than the fan which included complete stages. Another machine with complete stages should be calibrated to verify whether or not all machines yield an independent set of correlations. The solution for each machine was unique, and in the future, this process could be automated with an optimizer to gather the ADD LOSS and ADD DEVIATION data. The next step in this process is to implement the correlations and verify that the calibrated code will yield pressure ratios and efficiencies that closely match the experimental results.

REFERENCES

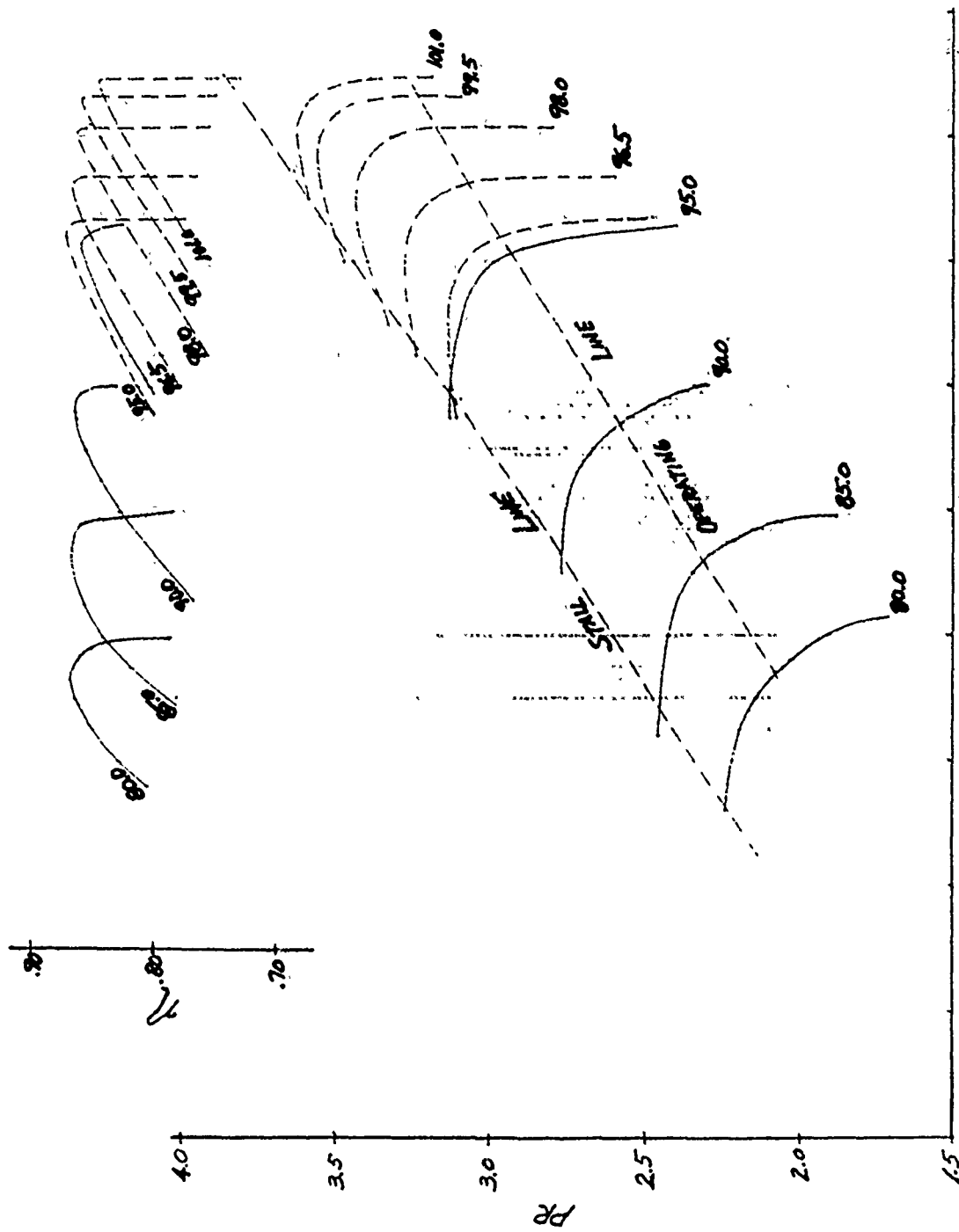
1. Krabacher, K.W. and J.P. Gostelow, "Single Stage Experimental Evaluation of High Mach Number Compressor Rotor Blading Part 4 - Performance of Rotor 2D," NASA CR-54584, October 6, 1967.
2. Fan data for COCODEC generated at Arnold Engineering Development Center, 1986.
3. Roberts, W.B., "COCODEC Code Enhancement - Project D185." Final Progress Report on Sverdrup Subcontract No. A9S-024, September, 1989.

APPENDIX
Performance Maps

Rotor 2D Performance Map



Three Stage Fan Performance Map



CORRECTED AIR FLOW U_{hm}/sec

DYNAMIC MODELING OF A DUAL FLOW PATH COMPRESSION SYSTEM

Jules W. Lindau V

ABSTRACT

A dynamic, stage-by-stage, post-stall compression system model intended for multiple flow path systems such as turbofans was investigated. Accurate results have been obtained with single flow path systems. However, the model exhibits numerical failure during deep surge in multiple flow path systems. Several deficiencies were identified which contribute to numerical failure: interpretation of stage characteristics at choked flow, an improperly imposed Mach number boundary condition, and a critical algorithm modeling flow at the splitter joining core and bypass flows to upstream flow. The stage characteristic interpretation and the improperly imposed boundary condition were corrected. The investigation suggested that the flow at the splitter region cannot be dynamically captured with a one-dimensional algorithm.

INTRODUCTION

DYNTECC, a dynamic, stage-by-stage, post-stall axial compression system model allows the detailed investigation of compression systems, like one finds in a fighter engine, during operation in the potentially dangerous post-stall regime. Currently DYNTECC, developed at Arnold Engineering Development Center, is used to model complete post-stall behavior in single flow path systems and full forward flow dynamic events in split flow path systems, such as found in a turbofan (See Figure 1). DYNTECC has been successfully used to model post-stall events in single flow path systems (See Figure 2). However, it is unable to capture dynamic events that include reverse flow, such as deep surge, in a split flow path system (See Figure 3).

Aircraft engine stall is a complex flow phenomenon caused by sustaining a high angle of attack on compressor blades. Depending on the system dynamics, engine stall will develop into either rotating stall or surge (See Figure 2). Rotating stall causes low efficiency in the compressor resulting in low or no thrust from the engine and high temperatures that may damage the turbomachinery. Compressor efficiencies of 20 percent or less are common during rotating stall. Rotating stall was often referred to as nonrecoverable stall in the past because the only way to return to useful, thrust providing, operation was to shut down and restart the engine. Although rotating stall involves complex, time dependent phenomena within the compressor, the mass flow into and pressure rise across the compressor are relatively constant during fully developed rotating stall. Surge is characterized by regular, 3 - 20 Hz cyclic flow through the compression system. Surge, unlike rotating stall, returns to a flow coefficient and pressure rise that is in the stable operating regime of the compression system at some point during the cycle. Deep surge is characterized by complete reversal of flow as well as full forward flow through the steady-state operating point during each cycle.

Although most high performance aircraft are currently equipped with control strategies that avoid engine stall, a better understanding of post-stall behavior will lead to the development of inherently rotating stall proof compressors, compressors that may surge but never develop rotating stall. This should be an improvement upon control systems that limit the engine's operational envelope. Traditional, steady-state analysis of the turbomachinery in an axial compression system cannot capture the transient behavior of post-

stall operation. Therefore a dynamic model is necessary. The development of compressor stall is a result of the separation of flow from the blades when the angle of air flow incident on the blades becomes too large. Flow separation is a complex phenomenon which is difficult to model explicitly in the configuration of a complete compression system model. Thus a specialized technique for modeling post-stall phenomena in axial compression systems has been developed. Detailed analysis of the compression system requires an accurate description of the compression system's overall flow through geometry as well as a discrete representation of each stage in the compression system. In a turbofan engine a detailed model should include the flow through the fan, core, core compressor and bypass ducting (See Figure 1). Thus a stage-by-stage post-stall model is required.

DISCUSSION

DYNTECC is based upon a computational fluid dynamics model of the quasi-one-dimensional Euler Equations coupled with additional source terms representing the force and work imparted by the compression stage and bleed mass flow. These modified Euler Equations are integrated in time explicitly. Each compressor stage is represented, in a one dimensional, steady-state sense by a pair of stage characteristics (See Figure 4). The stage characteristics give a relationship between a flow coefficient through the stage and a pressure rise and efficiency across the stage. This is a steady-state relationship and does not yield the dynamic behavior of the stage. The dynamic behavior of the stage is determined by assuming that the development of the stalled flow field in the compressor blade row involves a first order time response. The linearized solution to this time response is then applied to the steady-state pressure ratio given by the stage characteristic at a given flow coefficient. The dynamic force in a stage is then a function of the flow coefficient in the stage at that point in the time integration, the dynamic force at the previous time step, and a time constant chosen to model the first order time response of the blade row. See Davis [1] and Greitzer [2] for details.

The Euler Equation solver used in DYNTECC is strictly based on the characteristic eigenvalues of the governing equations. Therefore it falls into a class of computational fluid dynamic algorithms known as upwind schemes [3]. In computational fluid dynamics the practice is to divide the solution domain into many subdivisions, called control volumes. Within each control volume, the solution is assumed to take a form represented by a specified mathematical relation. In this case the control volume simply contains a constant value for each flow property that is actually the mean value for that volume. An upwind scheme assigns a direction to the information that may be passed from one control volume to another based upon the characteristic wave speeds of the governing equations. The upwind method has yielded robust and accurate computational fluid dynamics solvers and has made DYNTECC more robust than its non-upwind, central difference, predecessor, DYNMOD.

Currently, DYNTECC yields accurate solutions for post stall behavior of single flow path models (See Figure 2), but fails (goes numerically unstable) when modeling deep surge in multiple flow path

systems (See Figure 3). The objective set for the ten week period of research was to quantify the reasons for the model's failure and, if possible, correct the failure mode.

RESULTS

The approach taken with DYNTECC was based on the assumption that the computational fluid dynamics aspects of the code were correct. The flow solver itself is a modular part of the code and has been used in several other computer models successfully. Therefore, the parts of DYNTECC that differentiate the code from a standard fluid solver and make it a dynamic post-stall compression system model were investigated. Source terms dynamically represent the compression system stages. Time-dependent, boundary conditions are used to throttle the flow and drive the simulation into stall. And a flow splitting algorithm separates the flow into core and bypass streams.

A systematic investigation of the assumptions made in the construction of the DYNTECC model for multiple flow path systems was performed. We were able to identify three physical violations in the modeling process and one possible modeling deficiency that was not a strict violation of the governing physics but might affect our ability to capture the overall flow pattern. The first physical violation was a boundary condition that unreasonably held a choked condition in an imaginary nozzle. Another violation was the interpretation of stage characteristics at flow coefficients greater than the choking value for a particular stage. The third involved the dynamic modeling of the flow at the plane separating the single flow path upstream from the core and bypass flow paths downstream. In addition to these physical violations, it was proposed that modeling the flow in the fan with separate core and bypass flows might be an improvement on modeling the fan with a single flow path.

BOUNDARY CONDITION IMPROVEMENT

The flow in the bypass duct was assumed to exit through a constant area nozzle. In an engine this exit would lead to an afterburner/mixer or to ambient air. This nozzle was assumed to be choked. Therefore the modeler could specify a certain Mach number in the exit plane of the bypass duct (See Figure 5). This exit plane would, in the actual engine, be upstream of the actual exit plane of the nozzle. However as long as the nozzle remained choked, it is correct to assume that at a given position upstream of the nozzle, the Mach number remains fixed. This is because, for steady, isentropic flow, in a choked nozzle, Mach number at a position is determined by the ratio of area at that position to throat area. However in order to remain choked, the total pressure in the nozzle must be sufficiently greater than the back pressure. For isentropic

flow with constant ratio of specific heats equal to 1.4, the total pressure must be 1.89 times the back pressure to choke a nozzle.

In examining output from DYNTECC, it was observed that flow in the bypass duct had exceedingly low densities just before failure. It was proposed that the fixed Mach number at the exit of the bypass duct was incorrectly removing mass from the bypass volume during surge cycles. This proposal could be defended if total pressure in the bypass volume were dropping below 1.89 times the back pressure. A flag was placed in the DYNTECC code to report when total pressure at the exit plane of the bypass duct dropped below 1.89 times the back pressure. This condition occurred, and an alternative to fixing the Mach number at a specific value was necessary.

Fixing the Mach number at the exit plane implies that a converging, choked nozzle must exist downstream of the exit plane. Therefore an actual converging section, made up of additional control volumes was placed at the former bypass exit plane. By specifying a back pressure at the new exit plane, the effect of fixing the Mach number at the old exit plane is held, while the nozzle is choked (See Figure 5). When the total pressure drops below 1.89 times the back pressure, and the nozzle unchokes, the solution remains physically realistic. In addition to specifying a back pressure for forward flow, the boundary condition was modified to allow reverse flow at the exit plane. In reverse flow, the former exit plane becomes an inlet. This requires the specification of two flow properties. One obvious boundary condition sets total pressure equal to the back pressure. Specifying total temperature would complete the set. However, no clear choice for total temperature was available. The only choice that gives continuous properties at the boundary saves the value of total temperature at the time step when flow initially reverses, and specifies that value as the boundary condition. This boundary condition is implemented whenever the flow at the exit plane reverses.

CHOKED STAGE INTERPRETATION

In modeling the compression system stages, force is added to the momentum equation based on the stage pressure characteristics, and work is added to the energy equation based on the stage temperature characteristics. These characteristics may exist in forms other than the pressure and temperature combination, but any alternative combination is merely another expression of the same information. In high

performance compressors, the upper limit on flow is due to choking of the flow in the compressor blade passages. In the compressor, the axial flow Mach number will still be subsonic when the flow is choked in the blade passage. If the flow coefficient is cast in a form representing axial Mach number, the curve representing pressure coefficient versus flow coefficient and the curve representing temperature coefficient versus flow coefficient will have infinite slope at the choking value. In an actual compression system, flow coefficients above this choking value may not be obtained. In a computational fluid dynamics model of flow through the compression stages with exact representation of the compressor geometry, flow coefficients above the choking value would violate the governing equations. However in DYNTECC, there is no mechanism, linked to the governing equations, to choke the flow in the blade passage. The compressor blading is represented by a force and work term only and does not contribute to the geometric representation of the system. Therefore the flow is not limited by the choking area, and exceeding the flow coefficient that corresponds to the actual choking in the compressor does not violate the governing equations.

It was observed that during simulation of the compression system, flow coefficients were exceeding choking values for certain stages. They were utilized as if they represented valid flow conditions in the compression system. Pressure and temperature coefficients corresponding to these flow coefficients are physically meaningless. In DYNTECC, the stage characteristics are stored as curve fits relating pressure and temperature coefficient to flow coefficient. These curve fits do not allow an infinite slope, and therefore choking is not represented in the proper sense. However, the curves intersect the abscissa at a value of zero pressure or temperature coefficient with a very large (but not infinite) slope (See Figure 4). For this work, the intersection point for each stage was considered to be the choked value. These curve fits were only meant to be used for a limited range of flow coefficient. When a value of flow coefficient is used that is outside the range the curve fits were meant to be used for, greater than the choked value, the results are unpredictable. During simulation, performance that violates the second law of thermodynamics was given by curve fits in this unpredictable region. Therefore, DYNTECC's interpretation of these curve fits needed to be altered when the choking flow coefficient was exceeded. Several options were explored. (1) The mass flow could be limited to less than the stage choking value by means other than the governing equations (It was already pointed out that the governing equations cannot be utilized to limit the flow at this condition).

(2) The pressure and temperature ratios in each stage could be limited to 1.0 for values of mass flow exceeding the choking value in that stage. Or (3) the pressure and temperature ratios in every compression stage in the system could be limited to 1.0 whenever any stage exceeds the choking value. The first option was ruled out because it violates conservation laws. The second option and third option are similar; option (2) creates a slope discontinuous shut-off in the source term at the choked stage itself while option (3) shuts-off all source terms whenever any stage is choked. Option (3) was chosen in order to mimic the physics in a manner that does not violate conservation laws but still tends to apply the most severe restriction on the flow.

FLOW SPLITTER MODELING

The critical element in properly modeling the dynamic post-stall flow in a split flow path system such as a turbofan is capturing the dynamic flow in the region that splits the inlet flow into hub and bypass streams. In DYNTTECC, the flow is modeled as one-dimensional; it may only vary axially at any point in time. The algorithm used to split the flow into a core and bypass stream must conservatively join a single upstream one-dimensional flow to two downstream one-dimensional flows (See Figure 6). In DYNTTECC, nodes for the inlet, core, and bypass regions are each located at the flow splitter plane. In this plane conservation laws are obeyed. However, with three nodes, this leaves a number of unknowns greater than the number of equations. For example, in forward flow, three equations are available from conservation, two equations are available from the characteristic form of the equations in the upstream region, and a single characteristic form of the equations is available from each downstream region (Recall that DYNTTECC uses a characteristic form of the governing, conservation equations). This leaves nine unknowns and only seven equations. Thus two additional equations must be imposed. For full forward flow the correct steady-state is achieved when any two total conditions are required to be equal in the core and the bypass nodes. However, this condition should not be imposed on fully reversed flow or mixed direction flow (Flow when the velocity at the core and bypass nodes have opposite signs). For fully reversed flow the proper steady-state is achieved when static pressure is held equal in the core and bypass nodes. This is equivalent to a Kutta condition and ensures that a smooth streamline is attached to the division between the core and the bypass. It is uncertain if any condition may be imposed when mixed direction flow occurs. If the logic utilized in the

flow splitting algorithm changes the conditions abruptly, the flow properties may be discontinuous in such a way that violates the conservation laws. The governing equations are isentropic. However, if they change abruptly, multiple conditions in the splitter may increase or (worse) decrease entropy.

In an actual turbofan during a deep surge, the core flow may reverse while flow continues to go forward through the bypass. In fact the hub region of the fan may contain reversed flow while the tip region flow is forward. This is a flow regime DYNTTECC cannot capture when the flow splitter is placed downstream of the fan. It was suggested that this limitation could be the cause of DYNTTECC's numerical instability during deep surge. If the flow is split upstream of the fan such that the streamline separating actual steady-state core and bypass flow is used to separate the hub from the tip, correct steady state flow will be captured and the additional flow regime of reversed flow in the core and forward flow in the tip may also be captured. This new configuration was modeled.

Changing the location of the flow splitter did not avoid the failure mode in the splitting algorithm. Mixed direction flow still occurred at the splitter plane during the simulation. This lead to failure during deep surge wherever the splitter was placed within the original computational domain. However, if a more robust flow splitter algorithm is developed, the position of the flow splitter will be left up to the modeler.

CONCLUSIONS

Throughout the process of altering DYNTTECC, it was observed that, for the split flow path model, during a reversal of flow and afterwards, flow properties became unrealistic. In particular, flow coefficients in the compressor indicated choking at mass flow values that were an order of magnitude less than the steady-state choking values for those stages. DYNTTECC captures deep surge for a single flow path. The sole modeling difference between the split flow path and the single flow path model is the inclusion of the flow splitting plane. Therefore the failure of the model must be due to the algorithm used at the flow splitting plane.

DYNTTECC still fails to capture deep surge in a split flow path configuration. However some recommendations can be made with regard to further investigation of this problem.

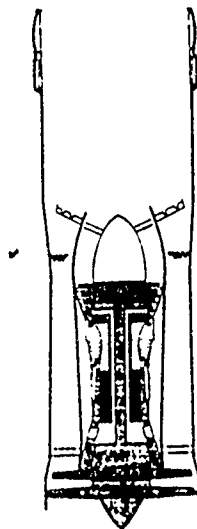
- (1) Until a more sophisticated method of choking the flow is developed continue to set all pressure and temperature ratios to 1.0 when a stage appears choked.
- (2) Until a better scheme for handling all flow regimes is developed, use the flow splitting algorithm that doesn't equate conditions in the core and bypass streams.
- (3) Employ the geometric nozzle with specified back pressure rather than a fixed Mach number boundary condition.
- (4) Future modeling should consider a two-dimensional flow model in the region of the splitter plane.
- (5) If two-dimensional modeling must be avoided, a parallel compressor model that separates core and bypass completely should be considered.

REFERENCES

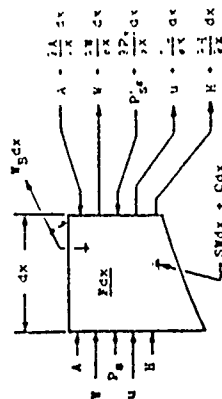
1. Davis, M.W., "A Post-Stall Compression System Modeling Technique," AEDC-TR-86-34, Arnold AFS, TN, February 1987.
2. Greitzer, E.M., "Surge and Rotating Stall in Axial Compressors, Part I: Theoretical Compression System Model," *ASME Journal of Engineering for Gas Turbines and Power*, V. 98, No. 2, pp190-198, 1976.
3. Yee, H.C., "A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods", NASA Technical Memorandum 101088, February 1989.

Dynamic Turbine Engine Modeling Technique

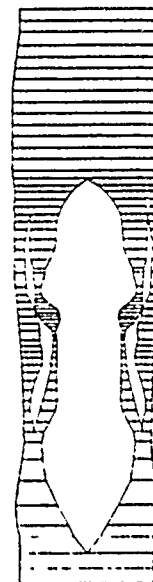
System Modeled



Elemental Control Volume



Control Volume Definition

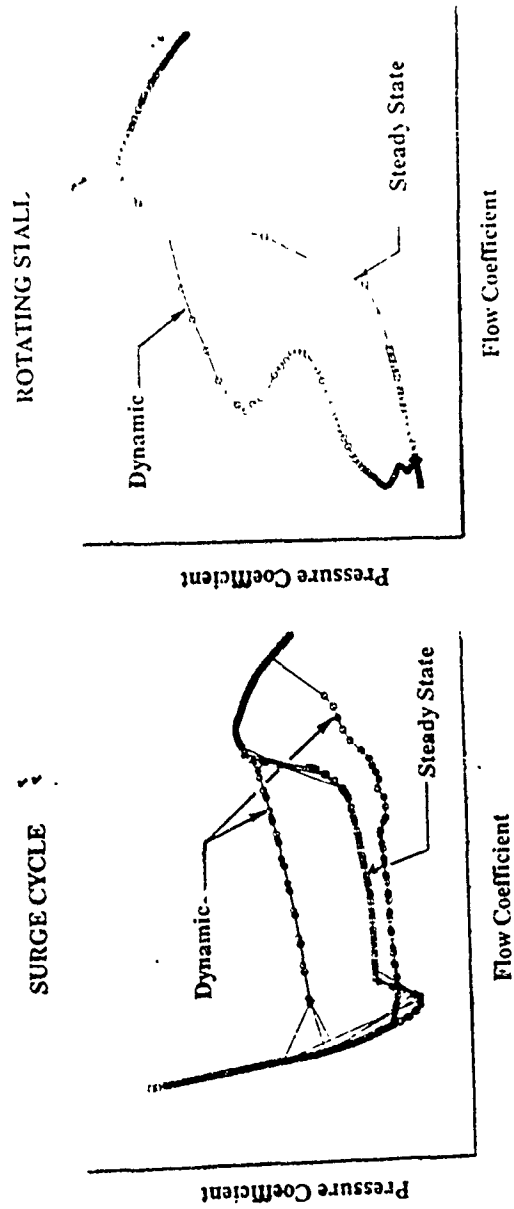


Modeling Technique

- 1-Dimensional
- Stage-by-Stage Maps
- CFD Flowfield Solution
- Continuity
- Momentum
- Energy

Figure 1. Dynamic turbine engine modeling technique.

Dynamic Compressor Characteristic During Post-Stall Events



- First Order Lagging Equation Applied in Rotating Stall Region to Provide Dynamic Blade Force

Figure 2. Dynamic compressor simulation during post-stall events.

Dual-Spool Compression System During HPC Surge Failure

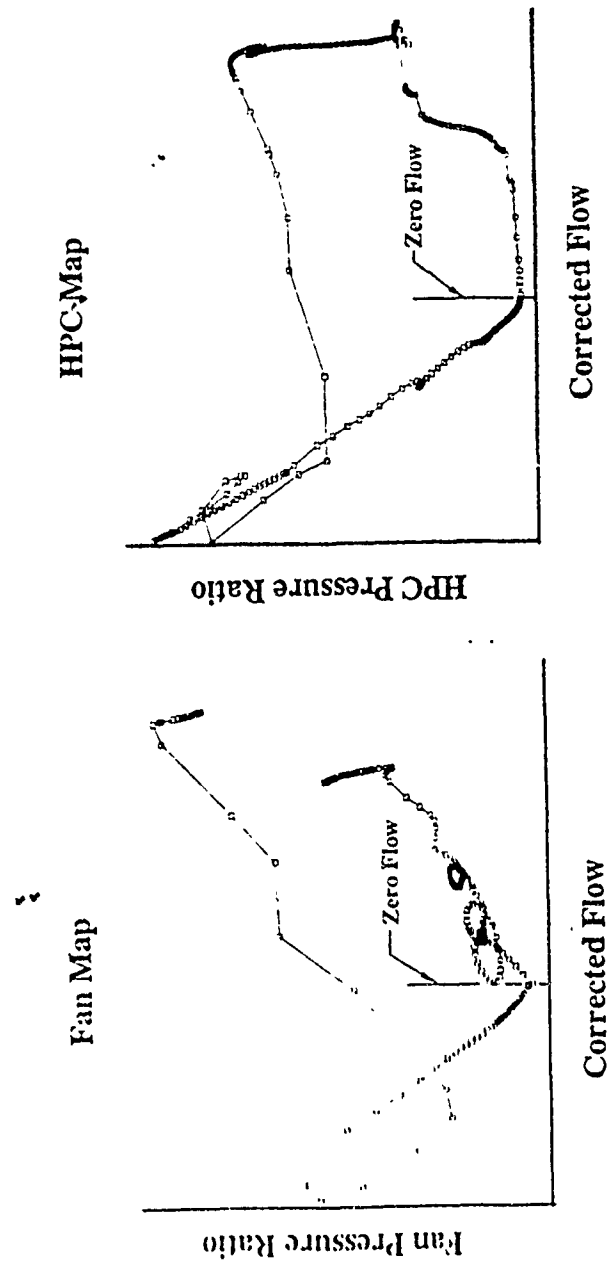
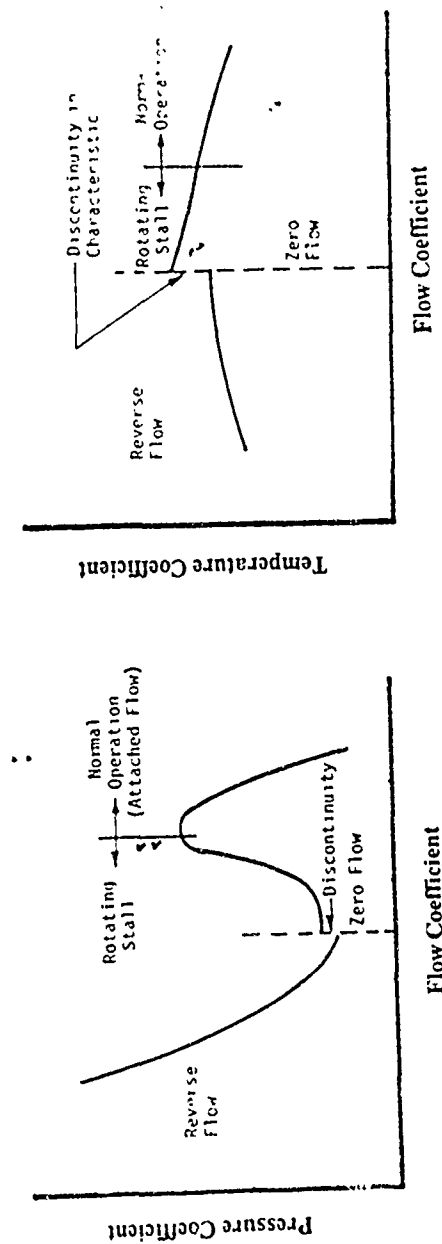


Figure 3. Failure mode of split flow path system.

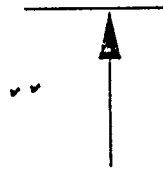
Typical Steady State Compressor Characteristics



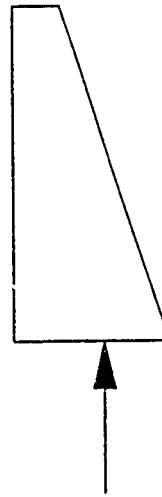
- Flow Coefficient Known from Last Time Step
- Stage PR & TR Calculated from Coefficients (Linear Interpolation Between Known Speedlines)
- SS Blade Forces & Shaft Work Calculated from PR, TR, & W
- First Order Lag Applied to Deliver Dynamic Blade Force in Rotating Stall Region

Figure 4. Typical steady state compressor characteristics

Nozzle Boundary Condition



Fixed Mach Number (subsonic)
requires $P_o/P_{back} > 1.89$

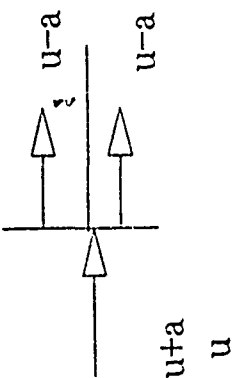


Nozzle geometry used. A/A^*
based on desired Mach number
at former exit plane.
Static pressure boundary condition
used based on characteristic.

Figure 5. Nozzle versus fixed mach number boundary.

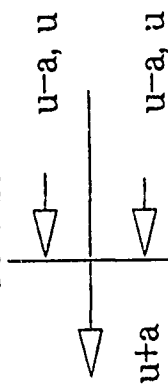
Flow splitter conditions (subsonic)

Forward flow



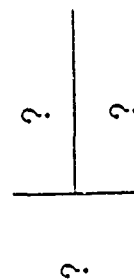
Best results occur when total properties (any two) are set equal in core and bypass.

Reversed flow



Best results should occur when static pressure is set equal in core and bypass.

Mixed flow conditions



No reason to link properties when flow is in opposing directions in core and bypass.

Figure 6. Conditions at flow splitter plane.

1991 USAF-UES SUMMER FACULTY RESEARCH PROGRAM
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the
Universal Energy Systems, Inc.

FINAL REPORT

Multigraph Implementation of Image Morphology

| | |
|--------------------|---|
| Prepared by: | Michael S. Moore |
| Academic Rank: | Graduate Student (M.S.E.E.) |
| Department and | Department of Electrical Engineering |
| University: | Vanderbilt University |
| Research Location: | Arnold Engineering and Development Center Arnold AFB, TN 37389 |
| USAF Researcher: | J.A. Nichols |
| Date: | 26 Sep 91 |

Multigraph Implementation of Image Morphology

by

Michael S. Moore

ABSTRACT

Morphological operators have been shown to be useful for image analysis and enhancement. Image morphology, though, is computationally expensive due to the amount of data, the complexity of data flow, and large required processing time. Multigraph, a system integration tool developed at Vanderbilt University, allows the building of complex algorithms from precoded modules (sub-algorithms). Multigraph will automatically schedule, execute, and control data flow between the steps of an algorithm. Multigraph is also capable of distributed processing using transputers, P.C.s, or workstations. This capability promises to make the morphological operations considerably faster. This report presents a multigraph implementation of a morphological algorithm for image enhancement. We will show that Alfa, a Multigraph graphics interface environment, allows the algorithm to be executed with ease and speed. Finally, we show a multigraph implementation of the algorithm that uses distributed processing between several workstations and give resulting speedup figures.

Acknowledgements

I would like to acknowledge the Air Force Office of Scientific Research for allowing me to participate in the 1991 Summer Research Program. The experience has enriched me both personally and academically.

Many people contributed to my project, and without them it could not have been as successful. I offer my thanks to Dr. Alan Peters, Ben Abbott, and Ted Bapty for their knowledgeable assistance and advice. I am especially grateful to Gabork Karsai for his infinite patience and hours of help.

I. INTRODUCTION

Morphological operators have proven to be effective in the analysis and enhancement of images. When large amounts of image data are to be processed, though, the implementation of the filters becomes computationally expensive. The images must be put through a series of processing steps, making data flow control a factor. Also, due to the inherent computational complexity of morphological operators, the required C.P.U. time is high. Ideally, we could implement the morphology in an environment that would allow us to specify the processing steps beforehand, and would automatically take care of the data flow between the steps. It is also desirable that this environment be capable of distributing the computational load between several processors. A tool is available that fits these needs. Multigraph. Multigraph was developed at Vanderbilt University to be used as a system integration tool for building complex algorithms from small precoded program modules. It schedules the execution of and controls the data flow between the modules, and is capable of distributed processing between several C.P.U.'s. This report describes the issues that were important to the implementation of a specific morphology algorithm, however, these issues can be considered to be applicable to any general multigraph application.

First we introduce the MultiGraph Kernel environment (MGK) and a Multigraph graphics interface (Alfa). Then, in light of our application, we consider data flow issues, buffers and data structures, and special considerations that must be addressed when using parallel processing. Lastly, we give the resulting speedup figures due to using distributed processing.

II. OBJECTIVES OF THE RESEARCH EFFORT

Implementations of image morphology algorithms are computationally expensive. This motivated us to consider Multigraph as a method of making morphological applications faster and more practical. My goals as a participant in the 1991 Summer Research Program were:

1. To produce general purpose image processing actors that read a Rasterfile into Multigraph memory, write a Rasterfile to disk, partition an image for parallel processing, etc...
2. To implement the "mclean" algorithm in the multigraph environment. I wanted the implementation to be easy to use and flexible. The mclean parameters needed to be readily modifiable to fit a specific application
3. To achieve an appreciable speed increase by using distributed processing.

As this paper will show, I was able to achieve these goals.

III. MGK: THE MULTIGRAPH KERNEL ENVIRONMENT:

The Multigraph Kernel (MGK) allows the programmer to construct a complex application from precoded program modules. It builds a "graph" representation of the algorithm, from which it schedules the processes and controls the data flow. The graph contains all the information about each process: its inputs, outputs, the code attached to it, and its execution environment. It also specifies the interconnections between the inputs and outputs of the processes (data flow). In the graph each process is represented by an "actor". (In this paper we will conveniently use the term actor for both the module of code that is executed by a process, and the process itself.) The inputs and outputs of an actors are called "ports". To build an application's graph in the MGK environment, the programmer must use the Kernel commands given in the Multigraph Users Manual (1) to create the actors and data nodes, specify interconnections, etc...

When distributed processing is desired, a version of MGK called netMGK must be used to build and execute the graph. The graph that represents our distributed morphology implementation is called `parallel_morph`, and the multigraph script "`morph cmd`" contains the commands that build it. As one can see from the `morph.cmd` script, building complex graphs directly in either of the MGK or netMGK environments can be complicated. For this reason, these environments are not optimal user interfaces for Multigraph. An easier, more ideal environment in which to build and execute application graphs would be one in which processes and data flow are represented pictorially. This ideal environment would automatically build the graph from this "schematic" of the algorithm.

IV. ALFA - THE GRAPHICS INTERFACE ENVIRONMENT:

Alfa, a Multigraph graphics interface environment, allows the programmer to fully automate the building of the graph. Using Alfa's graphics editor, icons are made for each actor. Each actor's icon represents a single process and is called a primitive. The primitives are the actual processing nodes where the actor scripts are executed. Primitives can be hierarchially interconnected with data nodes and other primitives to form compounds, and compounds can be combined to form higher level compounds (etc...). By interconnecting these primitives and compounds, a visual representation of

the algorithm's processes and data flow is constructed. When the system's pictorial model has been completed and stored as an icon, Alfa can be used to build the MGK graph and execute it. Presently, the Alfa graphics interface is not capable of building or executing distributed processes. However, we can still build distributed graphs with the help of an alternate version of Alfa, called AlfaCommand. AlfaCommand is designed so that instead of building the graph of the application automatically, it pipes the graph to a script called name.cmd (parallel_morph.cmd in our case). This script, if executed in MGK, will use only one processor. However, it can easily be modified with a text editor so that it will represent a distributed graph (such was the case for morph.cmd). This modified script can then be used to build and execute the distributed graph in the netMGK environment, which is fully capable of distributed processing. This inconvenience due to Alfa's lack of distributed processing capability is small when compared to the speed we expect to gain by distributing the processing load. Also, in the near future, we expect a new version of Alfa to be available that will be capable distributed processing.

V. BUFFERS AND DATA STRUCTURES IN MULTIGRAPH:

When a buffer is created by an actor, the normal allocation functions such as malloc and calloc can be used only if the buffer is local. It is our experience that these allocation functions should be avoided totally to avoid problems. Multigraph defines its own allocation functions which are much more predictable and safe to use. MGK provides two methods of buffer management. The command `mgk_autobuff_alloc(size)` allocates an automatically garbage collected buffer of length "size" that will be deallocated (freed) when it is no longer scheduled to be passed from any output port of any actor in the graph. (Note that Multigraph uses the term "propagate" instead of pass, so we will refer to an actor propagating a pointer instead of passing it hereafter.) Another allocation function is `mgk_buffer_alloc(size)`. The only difference is that this type of buffer is not automatically freed by the Kernel. It must be manually freed by using `mgk_free(buffer)`. In this application, we have found that the autobuffers are the better choice because of the automatic garbage collection feature prevents wasting of memory space.

It is very important to understand the way that Multigraph propagates buffers and data struc-

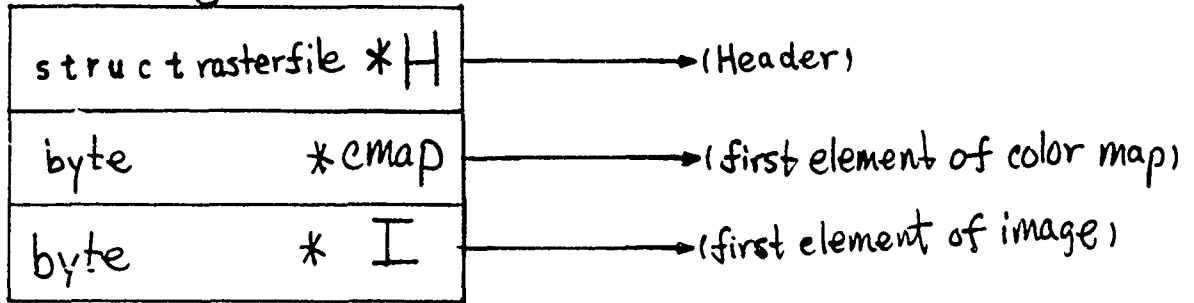
tures, so here we will describe our image data structure and its compatibility with Multigraph. In our actors, images in memory are referenced with the "RasImg" type data structure shown in Figure 1a. Notice that this structure contains three pointers:

1. "H" points to a Rasterfile type structure that contains the Raster Header information. The Rasterfile structure is shown in Figure 1b.
2. "cmap" is a pointer to the first byte color map array.
3. "I" is a pointer to the first pixel (byte) of the image array.

A pointer to this type of structure (* RasImg) can be propagated to actors that are in the same environment without ambiguity because the buffer addresses H, cmap, and I have meaning to the processor which allocated their buffer spaces. A problem occurs, though, when a RasImg type structure pointer is sent to a remote environment (which is the case when using distributed processes). When a structure (or buffer) pointer is passed to a remote process, MGK will automatically recognize it as a pointer and move its structure members (or buffer contents) to the remote processor's memory space. However, MGK will not recursively move the structure's members to the remote machine. (ie. It does not recognize structure members (or buffer data) as pointers and move their buffer contents). Thus, if a pointer to a hierarchical structure type such as RasImg is propagated, the remote actor receives a structure whose members are pointers to memory areas on another machine, and are thus meaningless. This fact was made painfully obvious by many hours of debugging. To avoid this problem, space for the Header, the ColorMap, and the Image is allocated in a single buffer which is called a RasBuff in the code (see the "RasBuffAlloc" subroutine). A pointer to a RasBuff can be sent to a remote machine with no ambiguity, since the buffer contains no pointers. After receiving the RasBuff, the actors construct the RasImg structure, which will only be used locally, by calling the "RasImg" subroutine. This construction of the RasImg structure using only the RasBuff is possible because all of the pointer offsets can be calculated from the Header values, which are always the first 8 integers in the RasBuff.

FIGURE 1: RasImg DATA STRUCTURE

a.) struct RasImg



b.) struct rasterfile

| | |
|-----|---------------|
| int | ras_magic |
| int | ras_width |
| int | ras_height |
| int | ras_depth |
| int | ras_length |
| int | ras_type |
| int | ras_maptype |
| int | ras_maplength |

V. FILE-MEMORY INTERFACE:

In this application, we used images stored in the Rasterfile format. In order to get the image data from the Rasterfile into multigraph memory, it was necessary to write an actor to opens the Rasterfile and reads its Header, ColorMap, and Image into a RasBuff. The actor OpenRas accomplishes this. OpenRas first opens the image file "File.Ras" (File.Ras is assumed to be under the directory specified in the environment variable RAS_IMAGES), then allocates a RasBuff of length determined from the Header as follows:

$$Length = length(Header) + length(cmap) + length(Image) \quad (1)$$

$$length(Header) = sizeof(struct rasterfile) \quad (2)$$

$$length(cmap) = H \rightarrow ras_maplength * sizeof(byte) \quad (3)$$

$$length(Image) = H \rightarrow ras_length * sizeof(byte), \quad (4)$$

where H is the Header pointer, and byte is as defined in include.h. OpenRas then propagates the RasBuff pointer from its output port.

The actor WriteRas is opposite in function to OpenRas. It receives a RasBuff pointer, constructs the corresponding RasImg structure, and writes it to a Rasterfile on the disk.

VI. MORPHOLOGY EXECUTION:

The morphology algorithm implemented in this work will be referred to as "mclean". The details of the mclean algorithm are not important to us here, but can be found in the Appendix. The mclean actor receives pointers to a RasImg type image structure and a parameter_file type structure. The latter structure contains the parameters that control the kind of processing mclean will perform, such as the size and type of structuring element, the number of partitions the image will be broken into, and the thresholds of the filter. The actor that allows the user to modify the mclean parameters is called "SetParameters". After performing all of the morphological filtering, mclean propagates the resulting enhanced image from its output port.

Using the four actors OpenRas, SetParameters, mclean, and WriteRas in a graph, the uniprocessor mclean implementation is complete. However, we are interested in using distributed processing, so we need further consideration of the problem.

VII. PARALLELISM CONSIDERATIONS:

Our approach to distributing the image processing load is to partition the image into pieces and perform the enhancement of each piece on a different remote machine. The enhanced partitions are then sent back to the main machine and used to construct the resulting enhanced image.

The partitioning is done by the actor "PartitionRas". PartitionRas receives a pointer to an input RasBuff, and breaks the image up into rectangular partitions, each of which is an independent RasImg type data structure (pointing to separate RasBuffs). Pointers to these image partitions are then propagated to the subsequent actors.

The partitions can be arranged in horizontal strips, vertical strips, or in windows, depending upon parameters M and N. The values of M and N are contained in the parameter_file structure and are defined as follows:

$$M = (\text{Number of vertical cuts} + 1) = (\text{Number of strips left to right}) \quad (5)$$

$$N = (\text{Number of horizontal cuts} + 1) = (\text{Number of strips top to bottom}) \quad (6)$$

$$\text{Total Number of partitions} = M * N \leq 9 \quad (7)$$

The reason that $M*N$ must be ≤ 9 has nothing to do with PartitionRas. It is merely because the graph parallel_morph only specifies 9 mclean actors. Theoretically, any number of partitions could be used if the graph were expanded. PartitionRas automatically selects the individual partition heights and widths so that the width (height) of the original image does not have to be divisible by M (N).

To prevent loss of image data due to parallelization, the partitions must overlap each other wherever a cut has been made in the original image. The width of the overlapping edges must be:

$$\text{vertical overlap width} = SEW - 1 \quad (8)$$

$$\text{horizontal overlap width} = SEH - 1 \quad (9)$$

where SEW is the Structuring Element Width, and SEH is the Structuring Element Height. Of course these overlapping edges will create some processing overhead, but it is expected to be insignificant compared to the total processing time. Note that if $M*N$ is not a prime number (4 6 8 9), using window partitions will cause less total overlap area than using strip partitions. Windowing $M*N$ partitions requires $(M-1)*(N-1)$ cuts, while $M*N$ strip partitions require $(M*N)-1$ cuts.

$$(M - 1) * (N - 1) = M * N - (M + N) + 1 \leq (M * N) - 1 \quad (10)$$

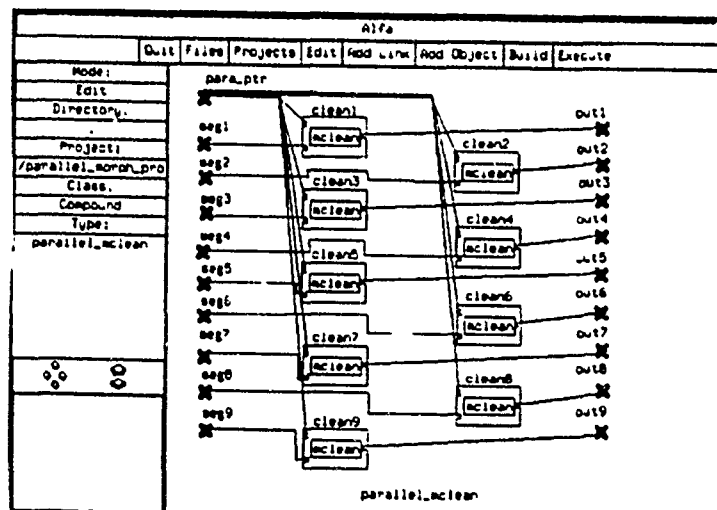
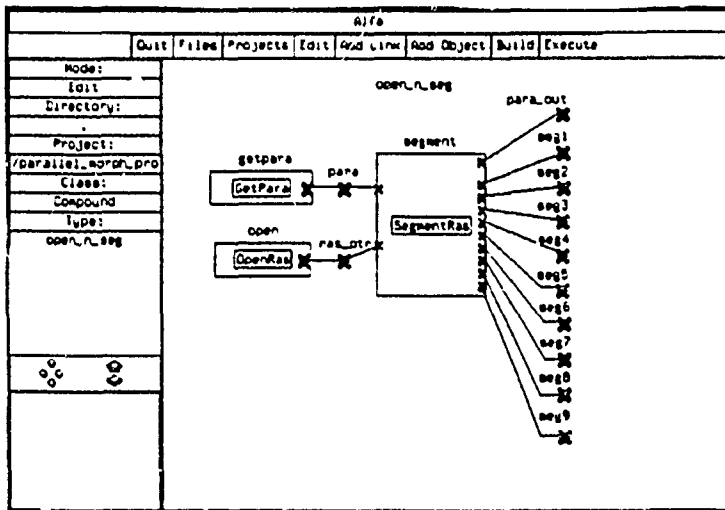
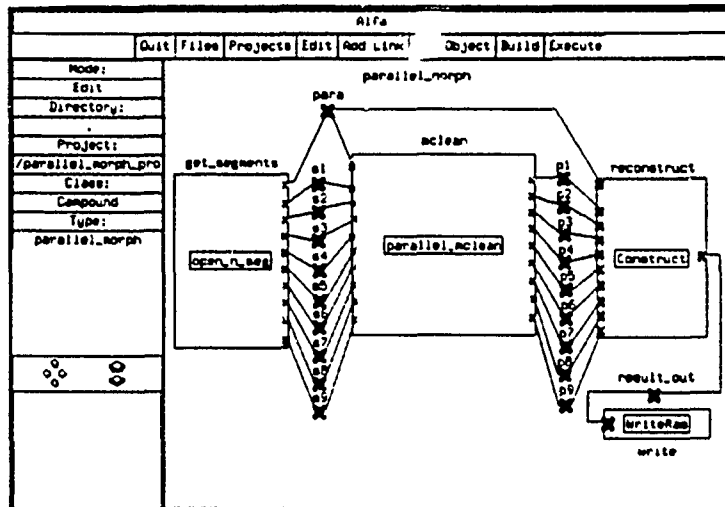
If $M*N$ is prime, of course, we have to use strip partitions.

The actor "ConstructRas" reverses this partitioning. It receives $M*N$ rectangular RasBuffs and constructs a new image from them that has the desired dimensions, taking into account their overlapping edges. It then propagates the resulting image from its output port.

When the system is run in parallel, each partition will be sent to a remote processor to be enhanced. In order to distribute the processing, then, there must be a separate mclean actor for every machine that is used. Each processors' mclean actor sends its enhanced partition to an input of ConstructRas, which is on the main machine. Refer to Figure 2 for further explanation.

In our parallel_morph graph, we allowed for up to nine remote processors, so the compound called parallel_mclean has nine separate mclean actors, nine image input ports, and nine image output ports. This graph with nine parallel mcleans can be used for any number of processors between 1 and 9. Only the first $M*N$ partitions will actually process any data (M is the number of vertical slices in the image when partitioning, and N is the number of horizontal, so the total number of partitions will be $M*N$). The other $9-(M*N)$ mcleans will do nothing. In order to achieve this flexibility, a trick was used. All of each actor's input/output ports must receive/propagate something, or the actor will never be triggered. When fewer than nine partitions are used, as is usually the case, PartitionRas propagates them from the 1_{ih} through $(N*M)_{ih}$ ports. From the rest of the ports a dummy RasImg structure is propagated which contains NULL as its header pointer. Mclean detects this dummy structure and immediately passes it to the next actor.

FIGURE 2: PARALLEL MCLEAN GRAPH



IX. SPEED GAINED BY DISTRIBUTED PROCESSING:

The parallel_morph graph was tested thoroughly using one, two, and three SUN4 sparc workstations. The speedup figures are quite encouraging. Since the mean processing time is dependent upon the structuring element size, data were taken varying not only the number of processors, but also varying S.E. size. Graphs of the processing time vs. S.E. size and processing time vs. Number of processors are given as Figures 3 and 4.

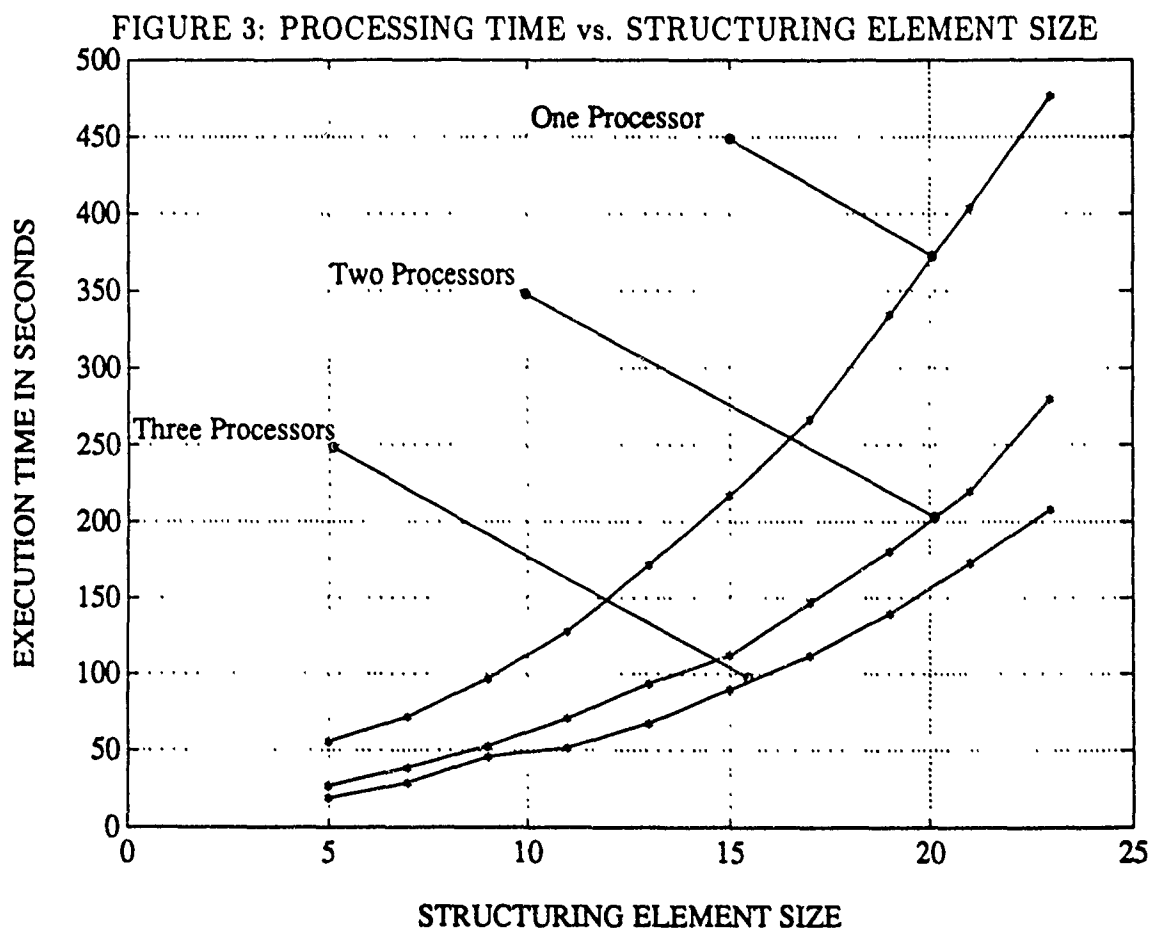
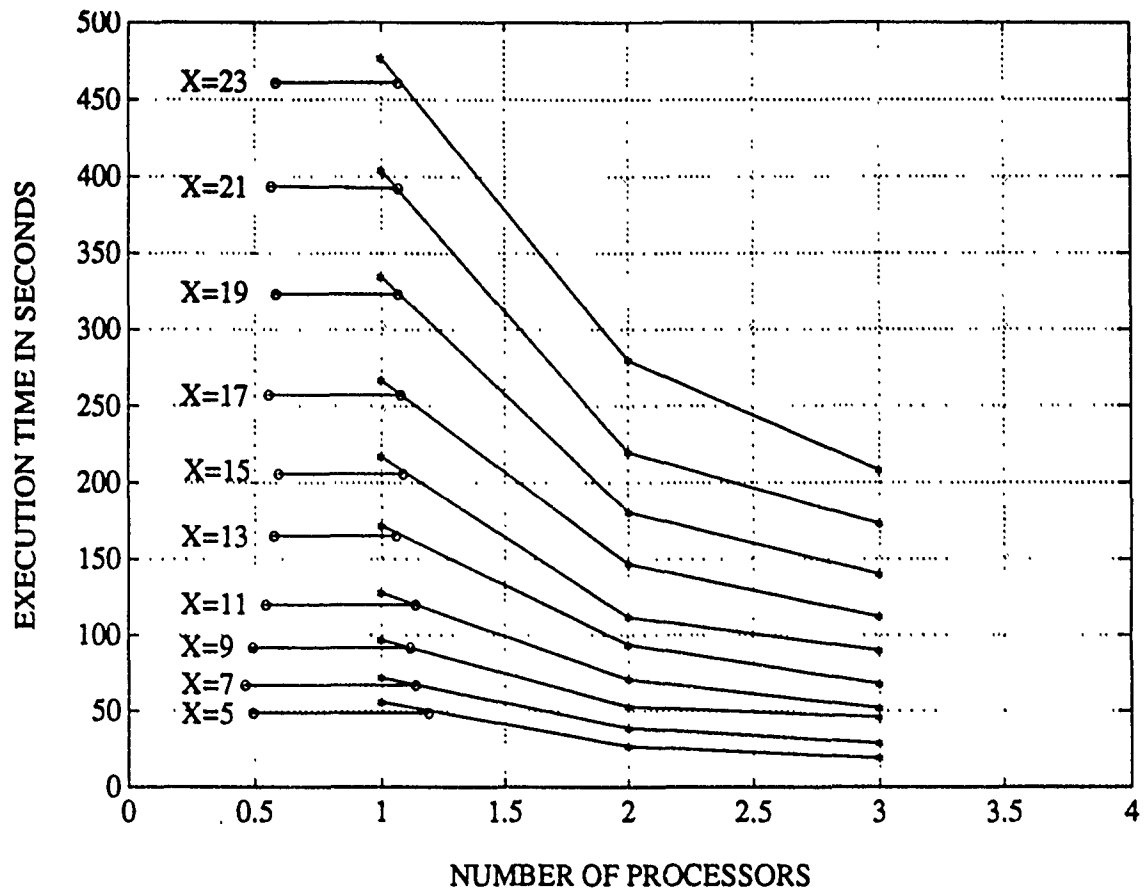


FIGURE 4: PROCESSING TIME vs. NUMBER OF PROCESSORS



VIII. RECOMMENDATIONS

Multigraph is a useful tool for implementating parallel image processing algorithms. In particular, we have shown that morphological filters can be made much faster and more useful by implementing them in the distributed Multigraph environment. This report showed one such implementation of a 2-dimensional morphological filtering algorithm.

A related topic that is of much interest, especially to Arnold Air Force Base engineers, is 3-dimensional morphology. 3-D spacetime morphological operators have been shown to be useful for reducing noise in time-sequence video images of rocket tests. 3-D morphology, though, is even more computationally expensive than the 2-D case. In 3-D the data flow control problems are magnified, and of course the morphology is more complex.

Because of these facts, it seems that it would be beneficial to implement 3-dimensional morphology operators in the distributed Multigraph environment. Then image-sequence algorithms could be more easily developed and implemented. This would potentially result in much faster and more effective implementations for the analysis of rocket test videos.

APPENDIX

MCLEAN ALGORITHM

The mclean algorithm uses the following morphological operations:

1. $OPENING(IMG) = Dialation(Erosion(IMG)) = Dialation \circ Erosion(IMG)$
2. $CLOSING(IMG) = Erosion(Dialation(IMG)) = Erosion \circ Dialation(IMG)$

where \circ denotes the composition operator.

The Dialation and Erosion operations correspond to approximation of the image from above and below, respectively, as the sum of shifted and scaled versions of the Structuring Element.

Steps of the Mclean Algorithm:

1. Take the Opening of the Image (OP).
2. Take the Closing of the Image (CL).
3. Form a smoothed version of the Image (SM) by averaging OP and CL.

$$SM = (OP + CL)/2 \quad (11)$$

4. Calculate the TopHat transform (Top) by taking the difference in the original image and the Smoothed version.

$$Top = IMG - SM \quad (12)$$

Presumably, the TopHat will contain all of the edge information and the noise present in the original image. Of course, we want to keep the edges and remove the noise.

5. Calculate the Histogram of SM, and calculate thresholds based on the Histogram. The mclean program uses the Histogram mean (+/-) σ_H as the (upper/lower) Thresholds. (σ_H is the standard deviation of the Histogram). Note we are assuming that the noise and small feature support of the TopHat transform are separable in the intensity domain.

6. Threshold the TopHat:

$$\widehat{Top} = Thresh(Top, Lthresh, Uthresh) \quad (13)$$

where the Thresh function sets the values of Top lying between Lthresh and Uthresh to zero, and sets the others to one.

7. Do a logical *and* of \widehat{Top} and Top to get $SSupp$; $SSupp$ (SmallSupport) will presumably contain only the support for the features in the image that are smaller than the structuring element.

$$SSupp = (\widehat{Top})and(Top) \quad (14)$$

8. Add the Small Support back into the Smoothed version of the image.

$$Enhanced\ Image = SM + SSupp \quad (15)$$

The Enhanced Image will contain most of the small features of the original, but will be relatively noise free.

1991 USAF-RDL SUMMER FACULTY RESEARCH PROGRAM
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Conducted by
Research and Development Laboratories, Inc.

FINAL REPORT

An Approximate Method for the Prediction of Underexpanded
Rectangular Nozzle Exhaust Plume Boundaries

| | |
|--------------------|---|
| Prepared by: | Kyle L. Nash |
| Academic Rank: | Graduate Student |
| Department and | Department of Aerospace Engineering |
| University: | The University of Alabama |
| Research Location: | Arnold Engineering Development Center Arnold AFB, TN 37389 |
| USAF Researcher: | Brian Wettlaufer Sverdrup Technology, Inc., AEDC Group |
| Date: | 1 August 1991 |

An Approximate Method for the Prediction of Underexpanded
Rectangular Nozzle Exhaust Plume Boundaries

by

Kyle L. Nash

ABSTRACT

An existing simple modeling technique is proposed for the prediction of underexpanded exhaust plume boundaries. This technique transforms axisymmetric plume boundary coordinates to the respective rectangular nozzle coordinates by using a simple model for the plume cross-section. The method employs a scaling factor to account for mixing rate differences between axisymmetric and rectangular nozzle exhaust plumes. Studies indicate that the technique shows promise as a preliminary design and analysis tool.

ACKNOWLEDGEMENTS

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsoring this research effort. I also wish to thank Research and Development Laboratories for their administration of this program which was greatly facilitated by the work of Carlos Tirres and Major David Hart at Arnold Engineering Development Center. Finally, I wish to express appreciation for the assistance received from Brian Wettlaufer and Terri Brock of Sverdrup Technology.

I. INTRODUCTION

A growing interest in thrust vectoring, airframe-engine integration, and stealth technology has led to the application of rectangular exhaust nozzles in existing and proposed aircraft. These rectangular or two-dimensional (2-D) nozzles emit plumes unlike those of axisymmetric nozzles. In order to perform analysis on this new family of nozzles, corresponding new analysis techniques must be developed. Generally, the more advanced techniques or programs such as Navier-Stokes codes are adaptable to wide ranges of applications. These tools, however, are cumbersome and expensive to operate. If only general information such as the location of the exhaust plume boundary is desired, then the use of such an advanced method is impractical. For this type of information, preliminary design tools are commonly used to provide approximate solutions. These preliminary design tools are often based on simplified models of complex flowfields.

Many approximate techniques based on flowfield models have been developed over the years for providing information about axisymmetric flowfields. The emergence of 2-D nozzles now requires a new family of preliminary design techniques. Such a technique, based on a simple model of a 2-D exhaust plume cross-section, is proposed for underexpanded rectangular nozzle exhaust plume boundary prediction.

II. OBJECTIVES OF THE RESEARCH EFFORT

To provide exhaust plume boundary approximations for underexpanded rectangular nozzles, a simple modeling technique is required. Because of the amount of existing data for axisymmetric nozzles, as well as the existing approximate methods for axisymmetric plume boundary predictions, a 2-D plume method based on a transformation of axisymmetric plume data is desirable.

Simple methods have been developed for predicting exhaust plume boundaries of underexpanded axisymmetric nozzles in both a quiescent atmosphere and an external flow (References 4 and 5). These methods efficiently provide good approximations of the initial portions of the jet boundary. The 2-D transformation technique to be used in conjunction with these methods should be comparable in simplicity and accuracy. In short, a single tool comprised of an axisymmetric plume prediction method and the transformation technique is required to quickly produce rectangular nozzle plume boundary approximations.

III. TRANSFORMATION METHOD

A general transformation technique was originally proposed by Chu, Der, and Wun (References 1 and 2) as a simple method for the prediction of temperature contours for infra-red (IR) signature analysis. It is reasonable to assume that the same transition of the plume and the same scaling factors would apply for plume boundaries as for

temperature contours. With this assumption, the aforementioned method is adopted for plume boundary transformations. Reference 1 provides a complete explanation of the method. A brief description is included here.

It is known that the plume emitted from a rectangular nozzle will become axisymmetric far downstream of the nozzle exit plane. Based on this knowledge, the plume model given in figure 1 is constructed. At the nozzle exit, the plume is assumed to conform to the nozzle cross-section. At some distance downstream, X_a , the plume becomes axisymmetric. At intermediate locations, a cross-section of the plume is comprised of semi-circular sides attached to a rectangular mid-section. Proceeding downstream, the rectangular section decreases in width until only the semi-circular sides remain at X_a . The dimensions of the semi-circular sections are determined from scaling of the corresponding axisymmetric plume dimensions. The scaling factor is designed to account for mixing differences between axisymmetric and rectangular nozzle plumes. Once the semi-circular side dimensions are known, the rectangular mid-section width can be computed using conservation of mass.

A simple formula is derived in reference 1 for X_a , where the plume becomes axisymmetric. In terms of aspect ratio, where $AR = w/d \geq 1$,

$$\frac{X_a}{d} = C \left(\frac{AR}{1+AR} \right) AR \quad (1)$$

Based on analysis of empirical data, C is given a value of 12.

The model requires determination of an axisymmetric plume for a nozzle of exit diameter, d, where d is also the height of the rectangular nozzle (Figure 2). After being scaled, this plume provides the boundary coordinates of the nonaxisymmetric plume. Reference 1 defines the scaling factor f "as the multiplier which, when applied to the radial and axial coordinates of a data point of the basic axisymmetric plume, gives the corresponding location of that point in the semi-circular side sections of the nonaxisymmetric plume."

A cubic spline can be derived for the variation of the scaling factor f with the axial distance X. This equation is

$$f = 1 + f_1 X + (3f_a - 2f_1 X_a - 3)(X/X_a)^2 + (2 + f_1 X_a - 2f_a)(X/X_a)^3 \quad (2)$$

where $f_1 = \tan(\xi)$

$$f_a = 2(AR/\pi)^{1/2}$$

and ξ = nozzle half-angle.

f_a provides the value of the scaling factor at X_a . Downstream of X_a , the plume remains axisymmetric and is directly scaled by f_a .

The final element to be determined is the width of the rectangular mid-section for a cross-sectional plane at a given downstream location (Figure 3). This width is

expressed as a function of the local scaling factor by the following equation:

$$\frac{w}{d} = AR - (\pi/4)f^2 \quad (3)$$

Once the axial location reaches X_q , $f = f_q$ and the width of the rectangular section becomes zero.

IV. RESULTS

The axisymmetric plume results used in conjunction with this transformation technique were generated by the program discussed in Reference 5. This approach assumes that the interaction between the nozzle flow and external flow, as well as the expansion of the plume, can be accounted for in the axisymmetric prediction code and then transformed with reasonable accuracy.

Relatively little rectangular nozzle plume data is currently available for comparison. Limited data was obtained from reference 3. This data represents a slightly underexpanded square nozzle. The nozzle exit Mach number and the external flow Mach number are 3. The ratio of the nozzle to freestream static pressure ratio is 2. Figure 4 is a comparison of plume cross-sections for reference 3 and the transformation method at $X/d = 4$.

The data is only available near the exit plane. Unfortunately, near the exit plane the transformation is least accurate since the model does not as accurately suit what actually occurs. Specifically, near the exit plane,

cross-sections will appear more as rectangular and will not appear to have semi-circular sides. Slightly downstream, where the model is more appropriate, accuracy should improve significantly. This hypothesis can only be tested by obtaining plume boundary data further downstream.

V. CONCLUSIONS

A simple method has been found for the approximate prediction of exhaust plume boundaries emitted from underexpanded rectangular nozzles. Some limited studies have been performed to validate the method. A lack of sufficient empirical data for a complete validation of the technique suggests the need to generate such data with a test program. More extensive studies need to be performed in order to determine the limits on Mach numbers, pressure ratio, aspect ratio, and other parameters for which the technique provides acceptable accuracy. The data used for the studies contained in this report demonstrate that the technique shows promise as a preliminary design tool. Further studies would serve to establish greater confidence in the method.

REFERENCES

1. Chu, C. W., Der, J., Wun, W., "A Simple 2D-Nozzle Plume for IR Analysis," AIAA-80-1808, AIAA Aircraft Systems Meeting, August 1980.
2. Chu, C. W., Der, J., "Modeling of 2D-Nozzle Plume for IR Signature Prediction Under Static Conditions," AIAA-81-1108, AIAA 16th Thermophysics Conference, June 1981.
3. Dash, S. M., Wolf, D. E., Sinha, N., "Parabolized Navier-Stokes Analysis of Three-Dimensional Supersonic and Subsonic Jet Mixing Problems," AIAA-84-1525, AIAA 17th Fluid Dynamics, Plasma Dynamics, and Lasers Conference, June 1984.
4. Latvala, E. K., "Spreading of Rocket Exhaust Jets at High Altitudes," AEDC-TR-59-11, June 1959.
5. Nash, K., Whitaker, K., Freeman, L. M., "A Simplified Exhaust Plume Prediction Method for Underexpanded Nozzles in Supersonic External Flows," AIAA-91-2371, AIAA/SAE/ASME/ASEE 27th Joint Propulsion Conference, June 1991.

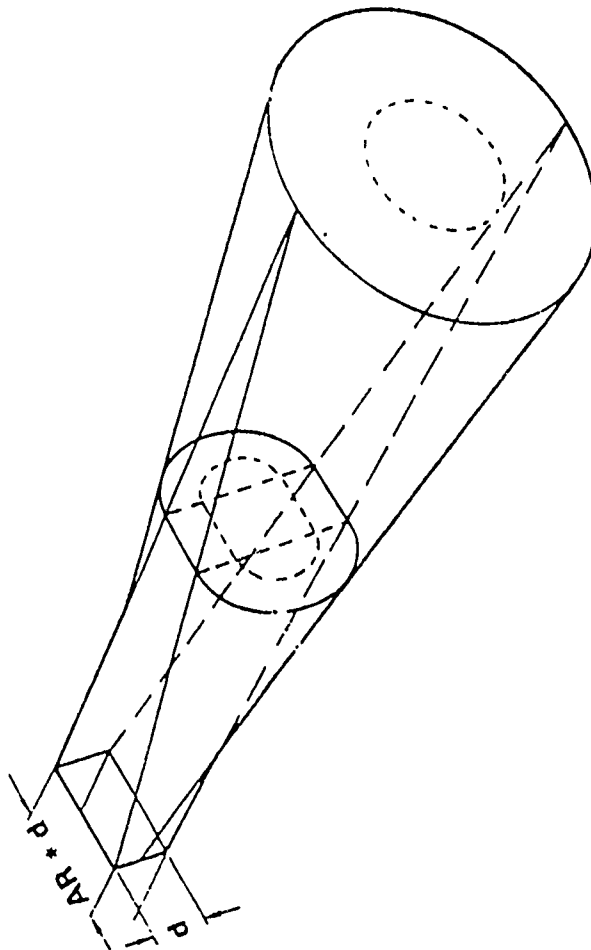


FIGURE 1: Transition of Exhaust Plume

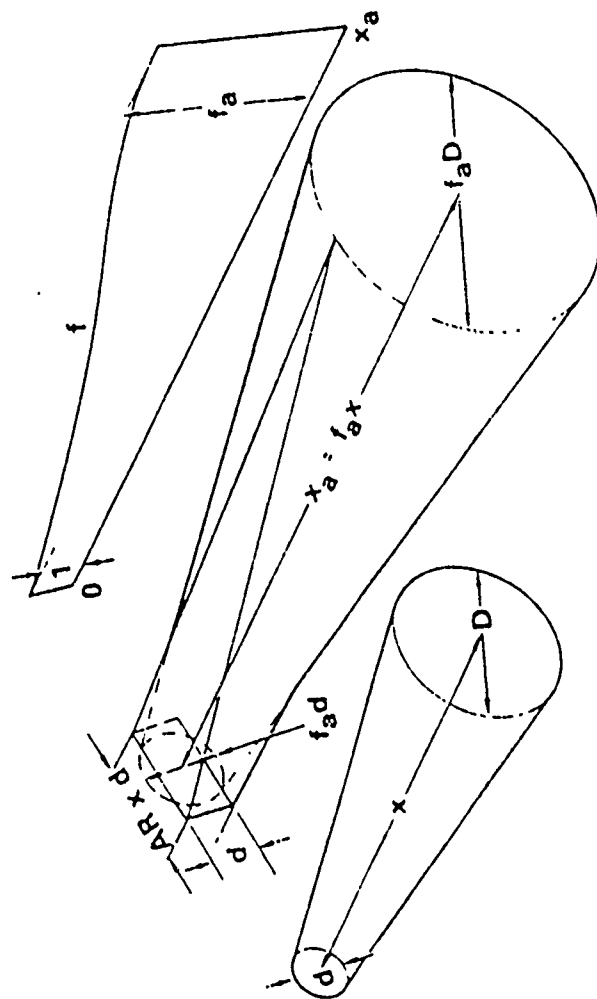


FIGURE 2: Scaling Factor Determination

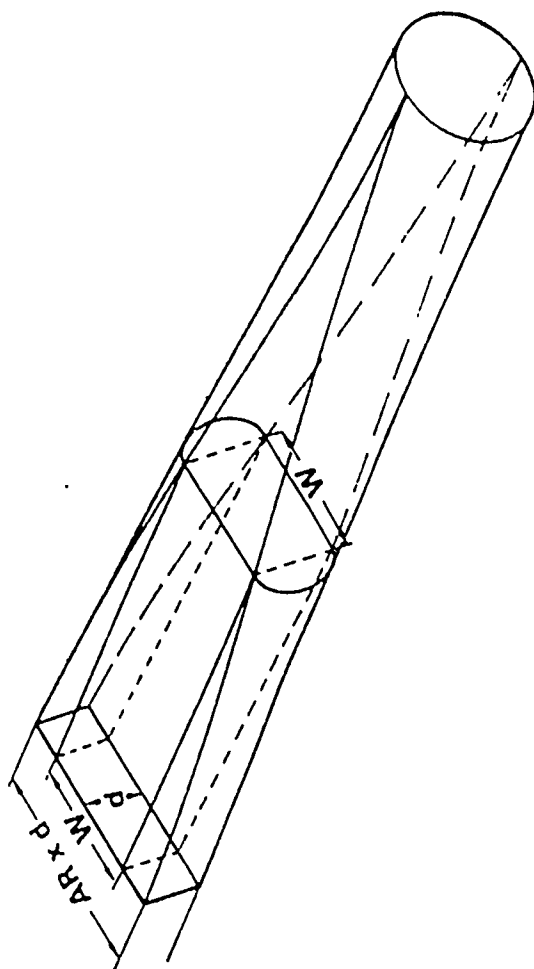


FIGURE 3: Exhaust Plume Width Determination

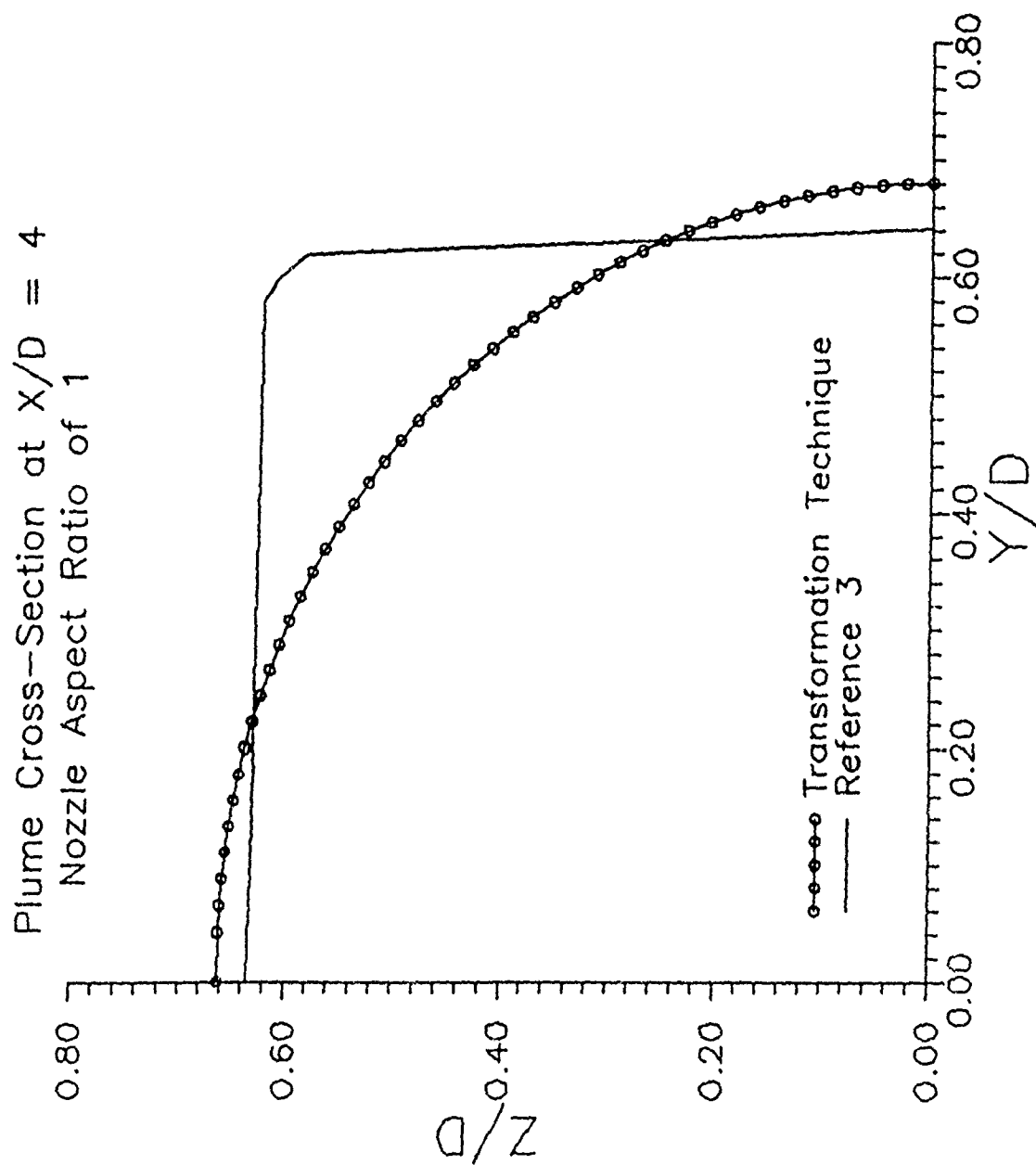


FIGURE 4: Comparison of Plume Cross-Sections

A Local Lagrangian Finite Volume Model for the Infinite Domain Shock Tube Problem

Blair H. Rollin

July 16, 1991

Abstract: An infinite domain shock tube model was constructed using the Local Lagrangian Finite Volume Method introduced to the author by K. C. Reddy ¹. For comparison purposes, two sets of initial conditions were chosen as test cases. These were the cases which Hirsch ² used as standard tests in order to compare the performance of various schemes. Several parameters were changed, and modifications were made to the basic scheme. The results indicate that the scheme has a large amount of numerical diffusion as well as a Courant-Friedrichs-Lewy (CFL) number restriction. Nevertheless, the method produced results comparable to many popular first-order methods. With a minor modification, the results were even more favorable.

1 Introduction

The Euler equations describe the flow of a perfect gas. Assuming sufficient differentiability, the one-dimensional equations can be written as

$$\int_V \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x} \right) \circ (Q, E) dV = 0,$$

where Q is the vector of conservation variables, and E is the vector valued function of the variables density ρ , velocity u , and internal energy e ,

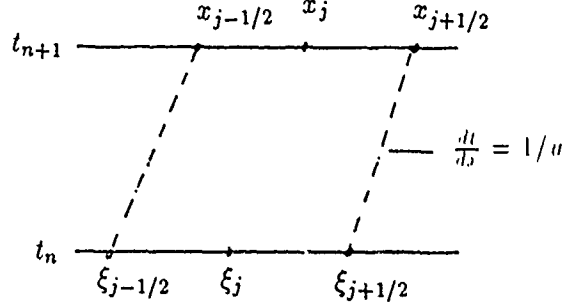
$$Q = \begin{bmatrix} \rho \\ \rho u \\ e \end{bmatrix}, \quad E = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ (e + p)u \end{bmatrix} \quad \text{with} \quad p = (\gamma - 1)(e - \rho u^2/2).$$

Or, by application of the divergence theorem, they can be written in the form of a surface integral about the space-time volume V with boundary ∂V ,

$$(1) \quad \oint_{\partial V} (Q, E) \circ n ds = 0.$$

The Local Lagrangian Finite Volume Method uses the form, (1), as a basis for discretization. For the one-dimensional problems, the discretization is performed on a line integral of one spatial dimension and time.

Consider a space-time discretized domain. We wish to evaluate the line integral (1) on the cell constructed between the lines of constant time, t_n and t_{n+1} , bounded spatially by the u -characteristics associated with the system drawn backward in time from $(x_{j+1/2}, t_{n+1})$ and $(x_{j-1/2}, t_{n+1})$ to the line t_n . These u -characteristics intersect the line t_n at $(\xi_{j+1/2}, t_n)$ and $(\xi_{j-1/2}, t_n)$ respectively.



We approximate these u -characteristics with straight lines of constant slope

$\frac{dt}{dx} = 1/((u(\xi_{j\pm 1/2}, t_n) + u(x_{j\pm 1/2}, t_{n+1}))/2)$. The normal vectors defined by this construction are then constant on each side of the grid cell. They are $(-1, 0)$ at t_n , $(1, 0)$ at t_{n+1} , $(-\bar{u}, 1)/\sqrt{\bar{u}^2 + 1}$ on the right side of the cell and, $(\bar{u}, -1)/\sqrt{\bar{u}^2 + 1}$ on the left side of the cell, where the bar indicates the average value of the function or variable evaluated at $(x_{j\pm 1/2}, t_{n+1})$ and $(\xi_{j\pm 1/2}, t_n)$.

Define $\xi_j = (\xi_{j+1/2} + \xi_{j-1/2})/2$. The vectors Q , and consequently E since it can be written in

terms of the components of Q , are approximated by $\bar{Q}, \bar{E} |_{j+1/2}$ on the right side of the cell, and $\bar{Q}, \bar{E} |_{j-1/2}$ on the left side of the cell. At t_n and t_{n+1} , they are approximated by $Q, E(\xi_j, t_n)$ and $Q, E(x_j, t_{n+1})$ respectively. The discretization of (1) then assumes the form

$$(2) \quad Q_j^{n+1} \Delta x + (-\bar{Q}_{j+1/2} \bar{u}_{j+1/2} + \bar{E}_{j+1/2}) \Delta t + (\bar{Q}_{j-1/2} \bar{u}_{j-1/2} - \bar{E}_{j-1/2}) \Delta t - Q(\xi_j, t_n)(\Delta x - (\bar{u}_{j+1/2} - \bar{u}_{j-1/2}) \Delta t) = 0,$$

These grid cells and the discrete equation (2) now form the basis for the solution of the flow field using the Locally Implicit Method (LIM) described by Reddy and Benek ³.

We incorporate the LIM method as follows. If we assume $Q^{n+1} = Q^{(m)} + dQ$ where (m) indicates our current estimate of Q^{n+1} we can solve (2) in terms of $dQ |_{x_{j-1}, x_j, x_{j+1}}$ by ignoring all second and higher order terms. Now, assuming $dQ |_{x_{j-1}} = dQ |_{x_j} = dQ |_{x_{j+1}}$, yields a linear system for dQ . We start with an initial guess for $Q^{(0)}$ and solve for dQ on each cell, sweeping across the domain. Iteration is manifested by multiple sweeps.

The intersection of the approximations to the u -characteristics and the line t_n are given by $\xi_{j\pm 1/2} = x_{j\pm 1/2} - \bar{u}_{j\pm 1/2} \Delta t$. The standard version of the Local Lagrangian Finite Volume Method uses linear interpolation between the two nearest grid points to evaluate variables at $(\xi_{j\pm 1/2}, t_n)$. The method was then modified for a higher order interpolation. It can be shown ² that (1) can be written as a decoupled system of ordinary differential equations where;

$$\text{entropy is constant along } \frac{dx}{dt} = u = \lambda_1,$$

$$u + \frac{2}{\gamma-1}a \text{ is constant along } \frac{dx}{dt} = u + a = \lambda_2,$$

$$u - \frac{2}{\gamma-1}a \text{ is constant along } \frac{dx}{dt} = u - a = \lambda_3, \quad a = \sqrt{\gamma p / \rho}.$$

$$\text{And entropy} = \text{const.} \implies \frac{p}{\rho^\gamma} = \text{const.}. \text{ Let } \omega_1 = \frac{p}{\rho^\gamma}, \quad \omega_2 = u + \frac{2}{\gamma-1}a, \quad \omega_3 = u - \frac{2}{\gamma-1}a.$$

Since the right and left sides of our grid cell can be assumed to be a good approximation to the

u -characteristics of the system, we perform higher order upwind interpolation (H.O.U.I) on each ω_i determining the upwind direction by the sign of λ_i which is evaluated at our current estimate of $Q(x, t_{n+1})$ consistent with the LIM method. We use the two grid points bordering $\xi_{j\pm 1/2}$ plus the next upwind point to fit a second degree Lagrangian interpolating polynomial to the data and use this for interpolation. The system of $\omega_i = k_i \quad \forall i = 1, 2, 3$ can be solved exactly for $u, \rho, p(\xi_{j\pm 1/2}, t_n)$ which then yield $Q(\xi_{j\pm 1/2}, t_n)$.

In this paper we examine the results from varying the number of LIM sweeps, CFL number, grid density, and interpolation schemes.

2 Problem Discussion

The shock tube or Riemann problem is that which we consider. A long tube, separated in the middle by a diaphragm, is filled with a calorically perfect gas. The gas is at high pressure to the left of the diaphragm and low pressure to the right. The diaphragm is instantaneously removed and a shock wave propagates into the low pressure region while an expansion wave propagates into the high pressure region. We wish to calculate the flow properties at any point and time. The exact solution to this problem is known⁴, which makes it ideal for analysis of numerical schemes. All experiments were run with initial conditions (a) $p_L = 10^5; \rho_L = 1; p_R = 10^4; \rho_R = 0.125; u_L = u_R = 0$ and (b) $p_L = 10^5; \rho_L = 1; p_R = 10^3; \rho_R = 0.01; u_L = u_R = 0$ and run to times 6.1msec and 3.9msec respectively. The times shown were rounded to two decimal places. Because of length considerations, the results from case (b) were omitted from this paper. However, the results from case (b) and the conclusions drawn from those results were entirely similar to those of case (a).

3 Results

Figures 1 and 2 show results from two LIM sweeps vs. five LIM sweeps. It is clear that the LIM method converges rapidly in the context of the Local Lagrangian scheme and little or nothing is gained from a large number of sweeps. Also, there is no numerical diffusion associated with the sweeps. Results were produced using 15 sweeps that emphasized these conclusions.

Figures 2 and 3 show a comparison of results produced using different CFL numbers. Figure 3 represents twice as many time steps as figure 2. Although there is no explicit numerical diffusion added to the scheme, these results show a considerable amount of diffusion associated with the scheme; the amount of diffusion increasing with the number of time steps.

A comparison of figures 2 and 4 show the effects of grid refinement. The grid associated with figure 4 is twice as dense as that associated with figure 2. Both use five LIM sweeps. The resolution of the shock and contact discontinuity is much improved.

Figure 5 shows the results using the higher order upwind interpolation (H.O.U.I). Figure 5 represents results from two sweeps vs. five sweeps associated with figure 4. The (H.O.U.I) yields a much sharper shock and improves the resolution of the expansion fan considerably. There is also some improvement in the resolution of the contact discontinuity. Preliminary experiments indicate that higher order interpolation of all conservation variables from the u -characteristic side only, produces instabilities.

Previously, this method had been applied to viscous and inviscid Burger's Equations and found to have a CFL number restriction. So it was not surprising that the code would not run above a CFL number of 0.53 .

4 Conclusions/Summary

It has been demonstrated that the Local Lagrangian Finite Volume Method has the potential to perform well on fluid flow applications. Though very different from standard finite difference techniques, ideas from standard finite difference algorithms, such as upwind interpolation, may be combined with it. Shortcomings including CFL number restriction and numerical diffusion must be investigated further. It should be emphasized that this is an extremely preliminary examination whose main purpose is to demonstrate the viability of the scheme, which it has. It is hoped that the basic scheme will prove useful in preventing nonlinear instabilities without additional diffusion or upwinding.

References

- [1] Reddy K. C., The University of Tennessee Space Institute , Tullahoma Tennessee
- [2] Hirsch, C.(1990). *Numerical Computation of Internal and External Flows*, Chichester, West Sussex PO19 1UD, England: John Wiley and Sons Ltd.
- [3] Reddy, K. C. and Benek, J. A., "A Locally Implicit Scheme for 3-D Compressible Flows," AEDC-TMR-90-P10, September 1990
- [4] Liepmann, H. W. and Roshko, A.(1957). *Elements of Gas Dynamics* , New York: John Wiley and Sons Inc.

Shock Tube Flow: Exact vs Local Lagrangian Finite Volume Method

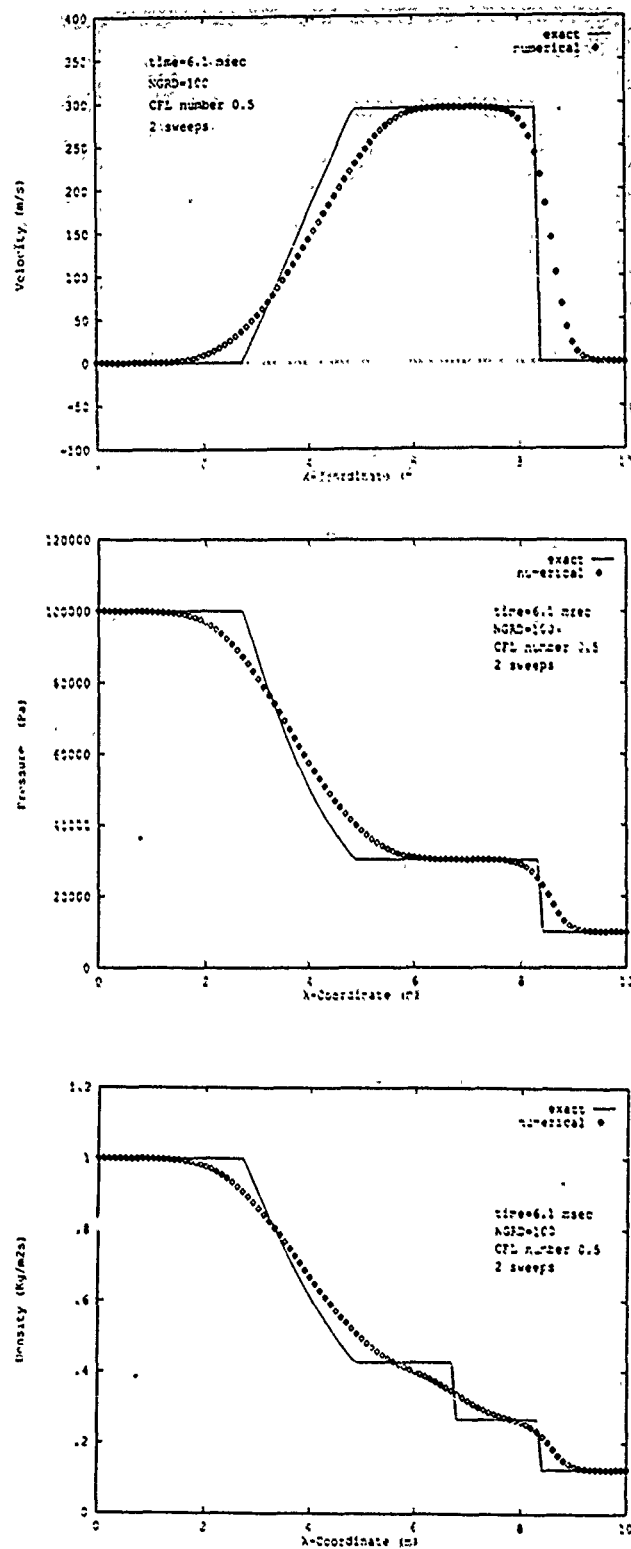


FIGURE 1: $p_L = 10^5$; $\rho_L = 1.0$; $p_R = 10^4$; $\rho_R = 0.125$; $u_L = u_R = 0$;
CFL=0.5; NGRD=100; 2 sweeps

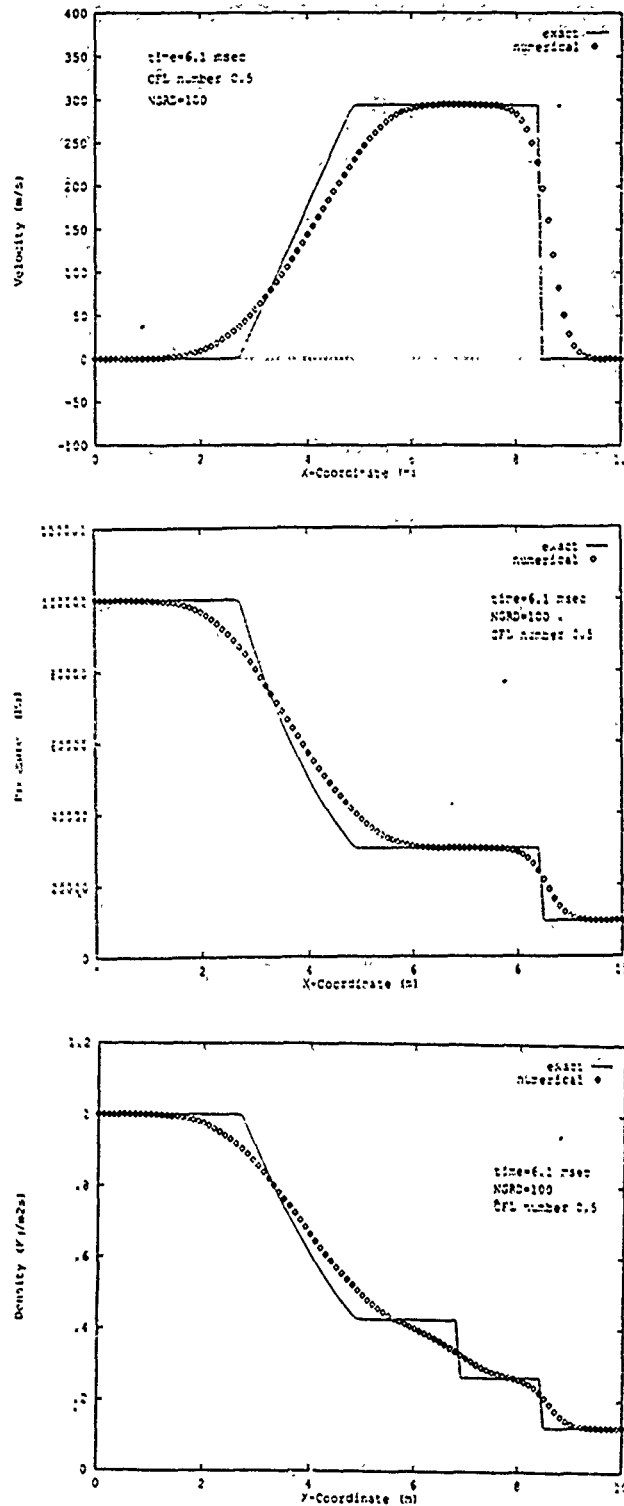


FIGURE 2: $p_L = 10^5$; $\rho_L = 1.0$; $p_R = 10^4$; $\rho_R = 0.125$; $u_L = u_R = 0$; CFL=0.5; NGRD=100; 5 sweeps

Shock Tube Flow: Exact vs. Local Lagrangian Finite Volume Method

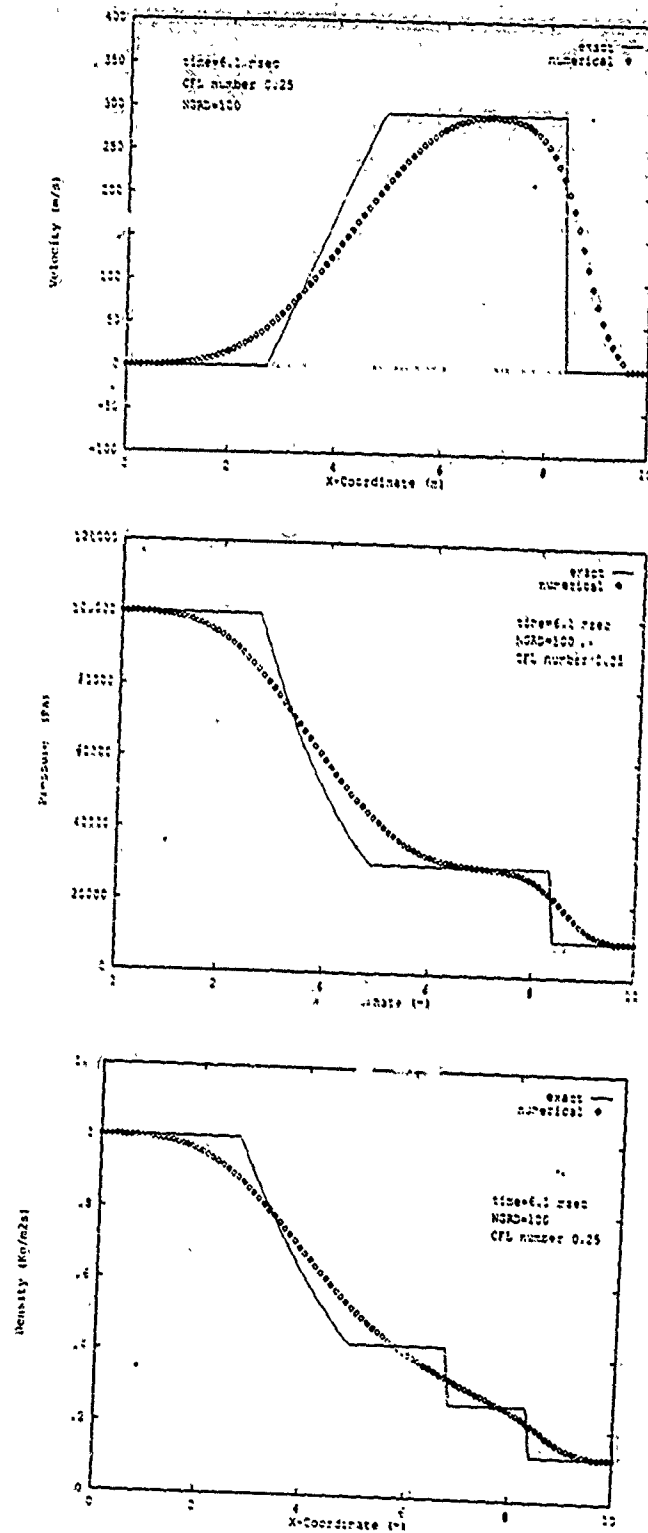


FIGURE 3: $p_L = 10^5$; $\rho_L = 1.0$; $p_R = 10^4$; $\rho_R = 0.125$; $u_L = u_R = 0$;
CFL=0.25; NGRD=100; 5 sweeps

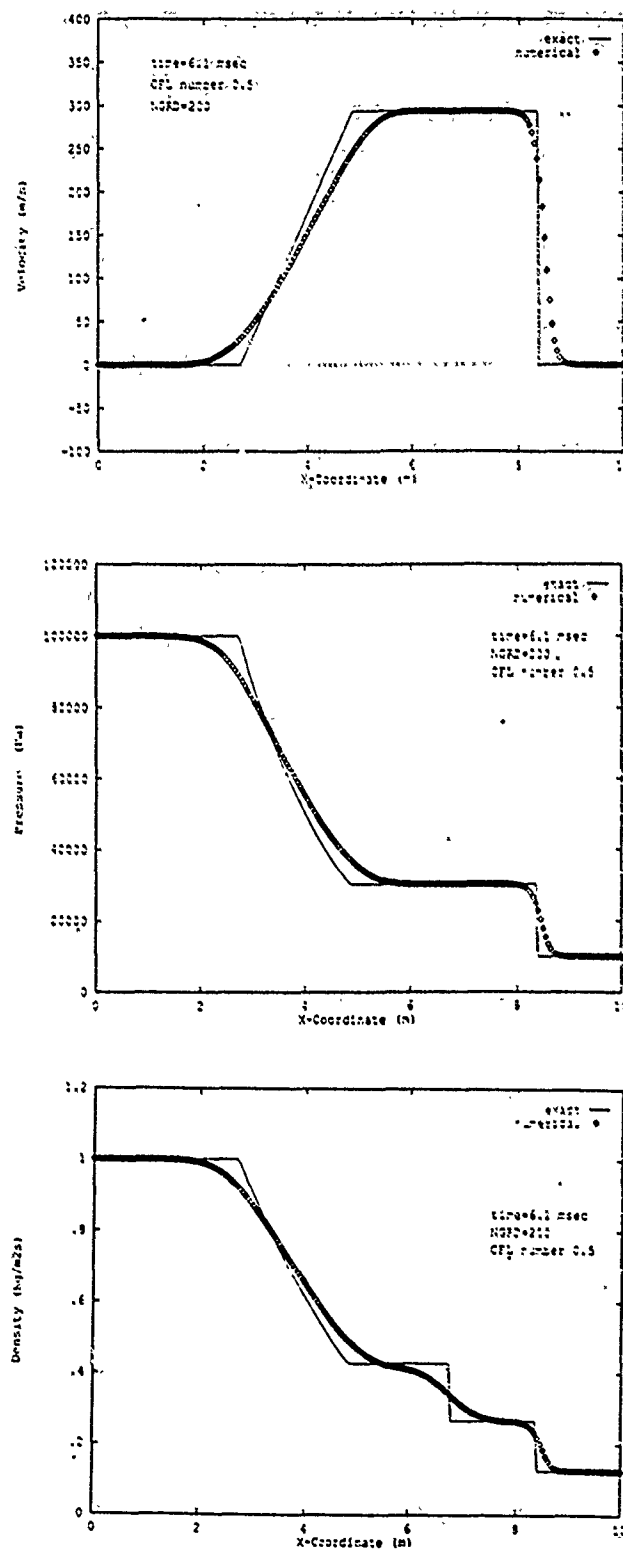


FIGURE 4: $p_L = 10^5$; $\rho_L = 1.0$; $p_R = 10^4$; $\rho_R = 0.125$; $u_L = u_R = 0$;
CFL=0.5; NGRD=200; 5 sweeps

Shock Tube Flow: Exact vs. Local Lagrangian Finite Volume Method

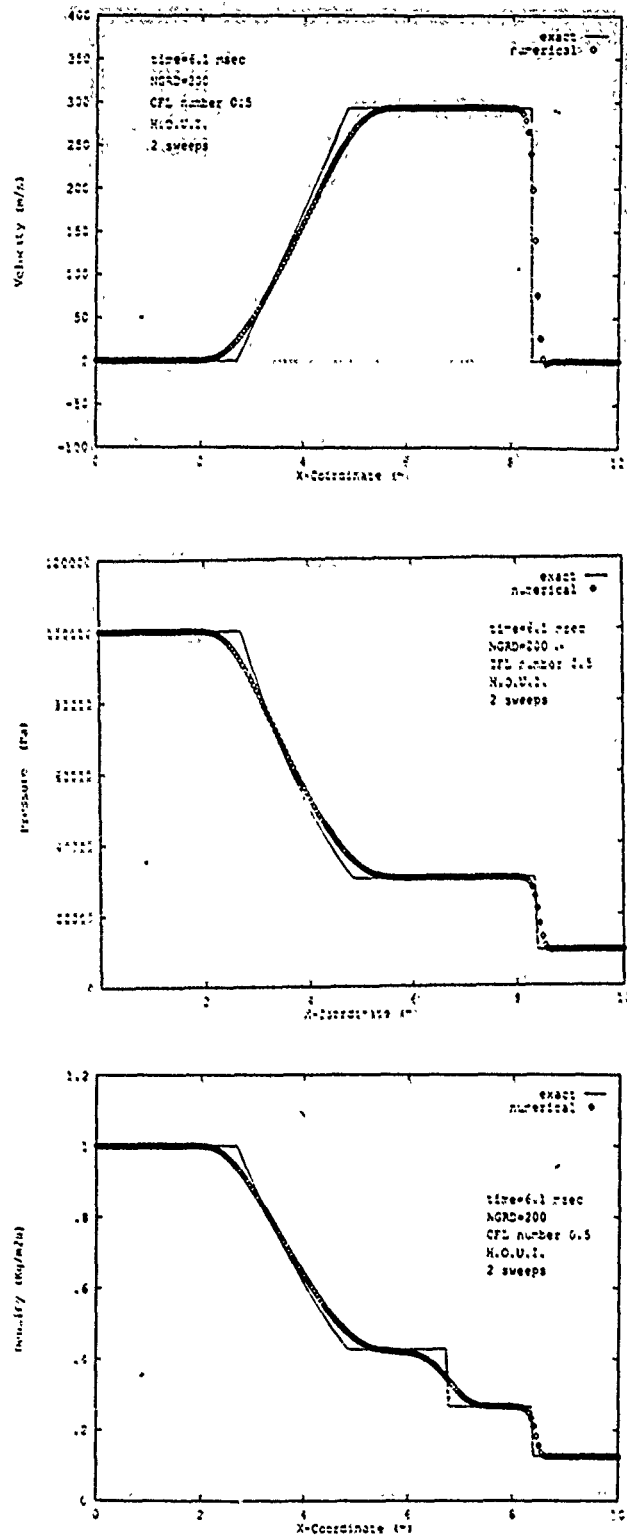


FIGURE 5: $p_L = 10^5$; $\rho_L = 1.0$; $p_R = 10^4$; $\rho_R = 0.125$; $u_L = u_R = 0$;
CFL=0.5; NGRD=200; 2 sweeps

COMPUTER MODEL FOR CAVITY OSCILLATIONS

Daniel Schatt

ABSTRACT

A computer model for flow over a cavity was developed to estimate the frequencies and magnitudes of the pressure oscillations that occur. The computer model can be used to estimate the maximum dynamic loads that will act on a store located in the cavity, and to help design small scale experiments that simulate full scale phenomena.

INTRODUCTION

Previous work has been done in three categories: experimental, Navier-Stokes solutions, and acoustical solutions. The most comprehensive experimental study, with and without stores, has been the WICS program at AEDC (Ref. 1). The experimental data obtained in this program are widely used to verify Navier-Stokes solutions and to help develop simpler, semi-empirical methods. Current Navier-Stokes codes are capable of predicting the complete flow phenomena, as shown by Dougherty in Ref. 2. However, computer time on the order of 100 cpu-hours of CRAY time is required. Suhs at AEDC (Ref. 3) has shown that it is possible to estimate mean pressure distributions and root-mean-square (rms) pressure distributions by Navier-Stokes solutions obtained in about twenty CRAY cpu-hours. However, since a computed spectrum is

not included, the results cannot be used to estimate the dynamic forces acting on a store located in the cavity. Simpler acoustical theories have been applied in an attempt to predict the spectrum, but these can only be used to predict the natural frequencies based on the dimensions of the cavity, and not the magnitude of the peak pressures that occur at these frequencies. In addition, experimental spectra show no peaks at these frequencies, but only at the so-called "edgetone" frequencies. In Ref. 4, Rossiter presents a simple theory for estimating the edgetone frequencies, which are produced by the shedding of vortices at the upstream edge of the cavity; these are, in turn, caused by acoustical waves emanating from the downstream edge. However, the magnitudes of the pressure pulses are not predicted.

The computer model was developed based on a theoretical concept that relates the edgetone frequencies to the acoustical frequencies and combines this with a method for estimating the maximum rms pressure. The details are presented in the following.

DISCUSSION

The turbulent mixing zone that separates the flow outside the cavity from that inside the cavity generates a continuous spectrum of sound waves that trigger the various edgetone frequencies that are observed experimentally. This implies that the edgetone frequencies should be used as the resonant frequencies in a frequency response equation.

Although the strengths of the sound waves may vary with frequency, the strengths (vorticity) of the created vortices are equal, since they are determined by the constant vorticity in the upstream corner. The sound waves are thought to act like a relay switch that triggers a much stronger pulse in the form of a vortex,

The damping ratio in the frequency response equation is crucial; it is determined by the relative magnitude of the edgetone and the resonant acoustical frequencies of the cavity. Minimum damping occurs when an edgetone frequency is equal to one of three possible resonant acoustical

frequencies, resulting in a maximum pressure amplitude at the edgetone frequency.

Experiment shows the maximum rms pressure in a cavity occurs at the downstream wall where the turbulent mixing zone impinges. Therefore, the maximum rms pressure in the cavity is equal to the rms pressure in the turbulent mixing zone.

RESULTS

A theoretical approach to predicting the fluctuating flow field in a cavity must include mathematical models of the following quantities:

1. Edgetone frequencies
2. Acoustical resonant frequencies
3. Maximum rms pressure in the turbulent mixing zone
4. Spectra reference pressure
5. Pressure on downstream wall of cavity as a function of time
6. Damping phenomena
7. Frequency response

The following mathematical models are not necessarily the best, but are sufficiently accurate to evaluate the overall concept.

1. Edgetone frequencies

The most widely used equation for estimating the edgetone frequencies was developed by Rossiter (*op. cit.*):

$$f_e = \frac{V_\infty (m - \gamma)}{L \left(M_\infty \frac{a_\infty}{a_t} + \frac{1}{\phi_d} \right)}, \quad (1)$$

where $m = 1, 2, 3, \dots$ = the frequency mode number of the edgetone.

Rossiter suggested two empirical constants in his original formula, γ and ϕ_d , which were shown to be essentially universal for deep cavities having thin initial boundary layers. The constant γ determines the distance beyond the downstream edge a vortex travels before an acoustical disturbance is generated that travels to the upstream edge and produces a new vortex. Experiment shows this distance to be 0.25 vortex wavelengths. The other constant, ϕ_d , represents the ratio of the average vortex velocity over the cavity to the freestream velocity. Experiment shows this constant to be 0.57 for thin initial boundary layers, but decreases as the boundary layer thickness increases, indicating that the vortices are moving at the dividing streamline velocity. In this analysis, the dividing streamline velocity ratio, including the effect of an initial boundary layer, is used. A rigorous equation could be used to calculate ϕ_d , such as the method presented in Ref. 5; however, the following semi-empirical equation is used here:

$$\phi_d = (0.0178 M_\infty + 0.6163) \left(1 - e^{-\frac{23}{\eta_p}}\right), \quad (2)$$

where $\eta_p = \sigma \frac{\delta}{L}$, a turbulent mixing position parameter,

and σ = similarity parameter for turbulent mixing.

For no bleed flow into the cavity, $\sigma = \sigma_0$, and,

for $M_\infty \leq 1$, $\sigma_0 = 12$,

and for $1 < M_\infty \leq 4$, $\sigma_0 = 8M_\infty + 4$,

and for $M_\infty > 4$, $\sigma_0 = 36$.

If there is bleed flow into the cavity, then the mixing is treated as two-stream mixing, and the similarity parameter can be determined from the following well-known equation:

$$\sigma = \sigma_0 \left(\frac{1 + \phi_c}{1 - \phi_c} \right), \quad \text{where } \phi_c = \frac{V_b}{V_\infty}.$$

The bleed flow velocity, V_b , is determined by assuming the bleed flow is injected uniformly over the upstream cavity wall at a density based on the freestream static pressure and total temperature. It is assumed that all of the bleed flow remains in the boundary layer, thus making it thicker. If it

is also assumed that the velocity profile remains unchanged, then the boundary layer thickness for conservation of mass flow is predicted by:

$$\delta = \delta_0 + \frac{\dot{m}_b}{W R_\infty V_\infty (1 - \frac{\delta^*}{\delta})}$$

where δ_0 = initial boundary layer thickness,
and, for no bleed, $\frac{\delta^*}{\delta}$ = ratio of boundary layer displacement thickness to total thickness.

The quantity $\frac{\delta^*}{\delta}$ is estimated using empirical equations devised by approximating the theoretical results presented in Ref. 6 with a 1/7-power velocity profile shape. In the subsonic and supersonic regimes, the results are:

for $M_\infty \leq 1$, $\frac{\delta^*}{\delta} = 0.0328 M_\infty + 0.125$,

and for $M_\infty > 1$, $\frac{\delta^*}{\delta} = 0.084 M_\infty + 0.0738$.

2. Acoustical resonant frequencies

There are three acoustical resonant frequencies corresponding to the length, width, and depth of the cavity. The fundamental length and width frequencies have wavelengths twice these dimensions. Since the wave velocity is the speed of sound inside the cavity where the static temperature is the total temperature of the freestream, the equations for these frequencies are the following:

Length: $f_L = \frac{a_t}{2 L}$ (3)

Width: $f_W = \frac{a_t}{2 W}$ (4)

The equation for the depth resonant frequency is more difficult to state because the top is open like

an organ pipe. The equation used was derived by R.C. Bauer at AEDC for a "tube" that corresponds to a deep cavity:

Depth:

$$fD = \frac{a_t}{2\pi D} \sqrt{\frac{2}{\gamma}} \quad (5)$$

It is important to note that these equations apply only to simple geometries at best, so other cavities--of unusual geometry or containing stores--will require a more elaborate acoustical analysis to describe the natural frequencies.

3. Maximum rms wall pressure

The maximum rms wall pressure occurs at the downstream edge of the cavity, and is produced by the impingement of the turbulent mixing zone at that point. An equation for estimating the rms pressure in a turbulent mixing zone was derived as follows, beginning with Bernoulli's equation:

$$dp + \rho u du = 0 ,$$

where ρ = static density in the free stream.

The rates of change of pressure, dp , and velocity, du , are treated as fluctuations attributable to turbulence. The rms of this equation is the following:

$$P_{rms} = \bar{\rho} \bar{u} u_{rms} ,$$

where $\bar{\rho}$ and \bar{u} are the mean values of density and velocity, respectively.

In the past, it was common in CFD codes to assume a linear relation between the Reynolds shear and the turbulent kinetic energy, so that

$$F_s = a \bar{\rho} (TKE) .$$

Define

$$C_f = \frac{F_s}{q_\infty} ,$$

where

C_f = friction coefficient,

and

q_∞ = freestream dynamic pressure.

Therefore,

$$P_{rms} = \bar{u} \sqrt{\frac{2 \bar{\rho} q_\infty C_f}{a}} \quad (6)$$

This equation has been applied successfully to a boundary layer, and so should also apply well to a turbulent mixing zone. For a turbulent mixing zone, the values of density, $\bar{\rho}$, velocity, \bar{u} , and friction coefficient, C_f , must be evaluated along the dividing streamline. The final equation for a mixing zone is

$$\frac{P_{rms}}{q_\infty} = \sqrt{\frac{4(1 - C_\infty^2) \phi_d^2 C_f}{a [1 - (C_\infty \phi_d)^2]}} \quad (7)$$

where

C_∞ = freestream Crocco number.

The remaining unknown in Eq. (7) is the coefficient of friction, C_f , along the dividing streamline. This can be determined by the method presented in Ref. 8, which applies the condition that "the momentum of the entrained mass flow must equal the total shear force along the dividing streamline." The resulting equation for no initial boundary layer is:

$$C_f = 2 \frac{(1 - C_\infty^2) I_d}{\sigma}$$

The normalized momentum of the entrained mass flow, I_d , was obtained by curve-fitting the theoretical values presented in Ref. 8. The equations are:

$$\text{For } M_\infty < 0.5, \quad I_d = 0.15,$$

$$\text{and for } M_\infty > 0.5, \quad I_d = 0.0338 M_\infty + 0.133.$$

The coefficient of friction, C_f , was corrected for the initial boundary layer by applying the well known experimental result that C_f in a fully developed mixing zone is ten times greater than the C_f in a corresponding boundary layer. The correction factor is in terms of η_p , the mixing position parameter defined in Ref. 9. The final equation for C_f is

$$C_f = 2 \frac{(1 - C_\infty^2) I_d}{\sigma} \left[0.9 \exp\left(\frac{-\eta_p}{8}\right) + 0.1 \right] \quad (8)$$

4. Spectra reference pressure

The absolute level of the pressure spectrum is determined by the strength of the vortices produced by the sound waves generated by the turbulent mixing zone. The reference pressure, P_{ref} , is the strength of these vortices, and is assumed to be constant for all frequencies. Since the overall rms pressure is determined by turbulent mixing, Eqs. (7) and (8), then the equation for the reference pressure is

$$P_{ref} = \frac{P_{rms}}{\sqrt{(0.5) \sum_{n=1}^{512} a_n^2}} \quad (9)$$

5. Wall pressure equation

The wall pressure equation is modeled as the sum of 512 sine waves of frequencies equal to the first 512 edgetone frequencies, and with (possibly) 512 different amplitudes, consistent with the 512 sets of Fourier coefficients determined from experimental data using the Fast Fourier Transform (FFT) method. However, the modeled equation is not exactly a Fourier series, since the difference in consecutive frequencies is not equal to the fundamental frequency; this is indicated by the phase term in the edgetone frequency equation (1). The pressure equation is:

$$\frac{P_{wall}}{P_{ref}} = \sum_{n=1}^{512} a_n \sin(\omega_n t) \quad (10)$$

where

$$\omega_n = 2\pi f_e \quad ,$$

and

P_{ref} = Vortex pressure strength (assumed constant for all frequencies).

The coefficient of each sine wave, a_n , is determined by the frequency response and damping ratio equations.

6. Damping ratio

Essentially, the damping ratio determines the relative magnitude of the pressure peaks in the spectra. For a very deep cavity, or tube, the damping ratio is determined by viscous effects and is given by the following equation:

$$d_{\mu} = \frac{8.885 \mu_t L W D a_t \sqrt{\frac{1}{\gamma}}}{P_{\infty} (L W)^2}, \quad (11)$$

where

μ_t = viscosity based on total temperature.

Applying this equation to a typical cavity that is relatively shallow compared to a tube yields an unrealistically small damping ratio when compared to experiment. In addition, since Eq. (11) is not a function of frequency all the pressure peaks in the spectrum are equal. Apparently another type of damping is occurring which is believed to be that predicted by the acoustical theories. Acoustical damping is attributable to the interaction of the various sound waves with an ultimate loss of energy out the open end of the cavity. This type of damping is referred to in this paper as "wave" damping, and is assumed to be a function of only the ratio of edgetone frequency to acoustical natural frequency. Based on experimental data from the WICS program, the following empirical equations for wave damping were developed for each of the first 10 modes, or edgetones:

Mode 1: $d_w = (0.006617 M_{\infty} + 0.003734) d_{w1} e^{d_{w1}}$

Mode 2: $d_w = (0.01284 M_{\infty} - 0.005529) d_{w1} e^{d_{w1}}$

Mode 3: $d_w = (0.006617 M_{\infty} + 0.0003734) d_{w1} e^{d_{w1}}$

Mode 4: $d_w = (2.837 M_{\infty} - 1.691) d_{w1} e^{d_{w1}}$

Mode 5: $d_w = (2.845 M_{\infty} - 1.7047) d_{w1} e^{d_{w1}}$

Modes 6-10: $d_w = (0.996 M_{\infty} - 0.5954) d_{w1} e^{d_{w1}}$

The total equation for the damping ratio is:

$$d = d_{\mu} + d_w \quad (12)$$

It is not certain that acoustical theories will predict wave damping since the acoustical-to-fluid mechanic coupling is not modeled, as can only be done through the complete Navier-Stokes equations.

7. Frequency response

In the experimental spectra from the WICS database, the minimum level is about the same for both low and high frequencies, quite unlike mechanical systems for which the frequency response continuously decreases for frequencies greater than the natural frequencies. Such response is, however, very similar to the frequency response characteristic derived by Bauer (Ref. 7) for unsteady flow in a tube. The general equation applied to a cavity is the following:

$$R_s = \frac{(1 + \frac{f}{f_e})^2 + 4(\frac{f}{f_e}d)^2}{(1 - \frac{f}{f_e})^2 + 4(\frac{f}{f_e}d)^2} \quad (13)$$

where

f = forcing frequency ,

f_e = edgetone frequency ,

and

d = damping ratio.

If a single forcing frequency is imposed on the cavity, then the response equation, Eq. (13), is the equation for each coefficient in the wall pressure equation. Various single forcing frequencies are used to obtain the theoretical spectra of the wall pressure. However, the actual wall pressure results from a continuous spectrum of forcing frequencies, so the forcing frequency in Eq. (13) is the corresponding edgetone frequency. As a result, the coefficients in the wall pressure equation, Eq. (10), are only a function of the damping ratio, as follows:

$$a_n = \frac{1 + d_n^2}{d_n^2} , \quad \text{where } n = 1, 2, 3, \dots, 512.$$

CONCLUSIONS

The computer program developed from the preceding model showed relatively good agreement with experimental data up to approximately Mach 1.5. At that point, the cavity acoustics began displaying higher frequencies and modes which the model could not predict. This indicates more work is required to better determine the complex interaction between fluid mechanic and acoustic phenomena. Nevertheless, the program does allow one to obtain fairly accurate predictions for a wide range of Mach numbers.

I have provided figures on the following pages which provide a visual comparison between the experimental results and the model prediction.

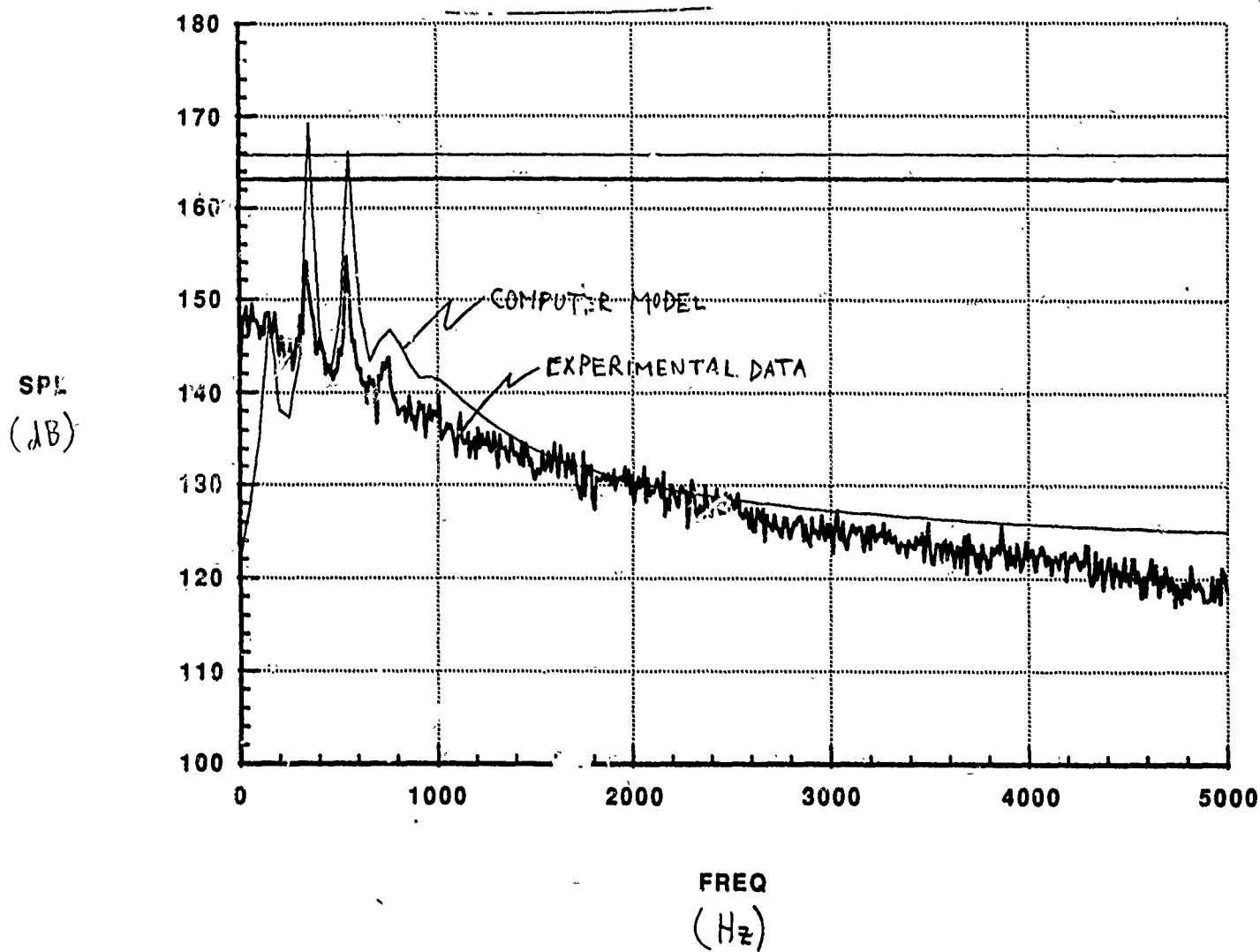


Figure 1. Mach No.=0.6

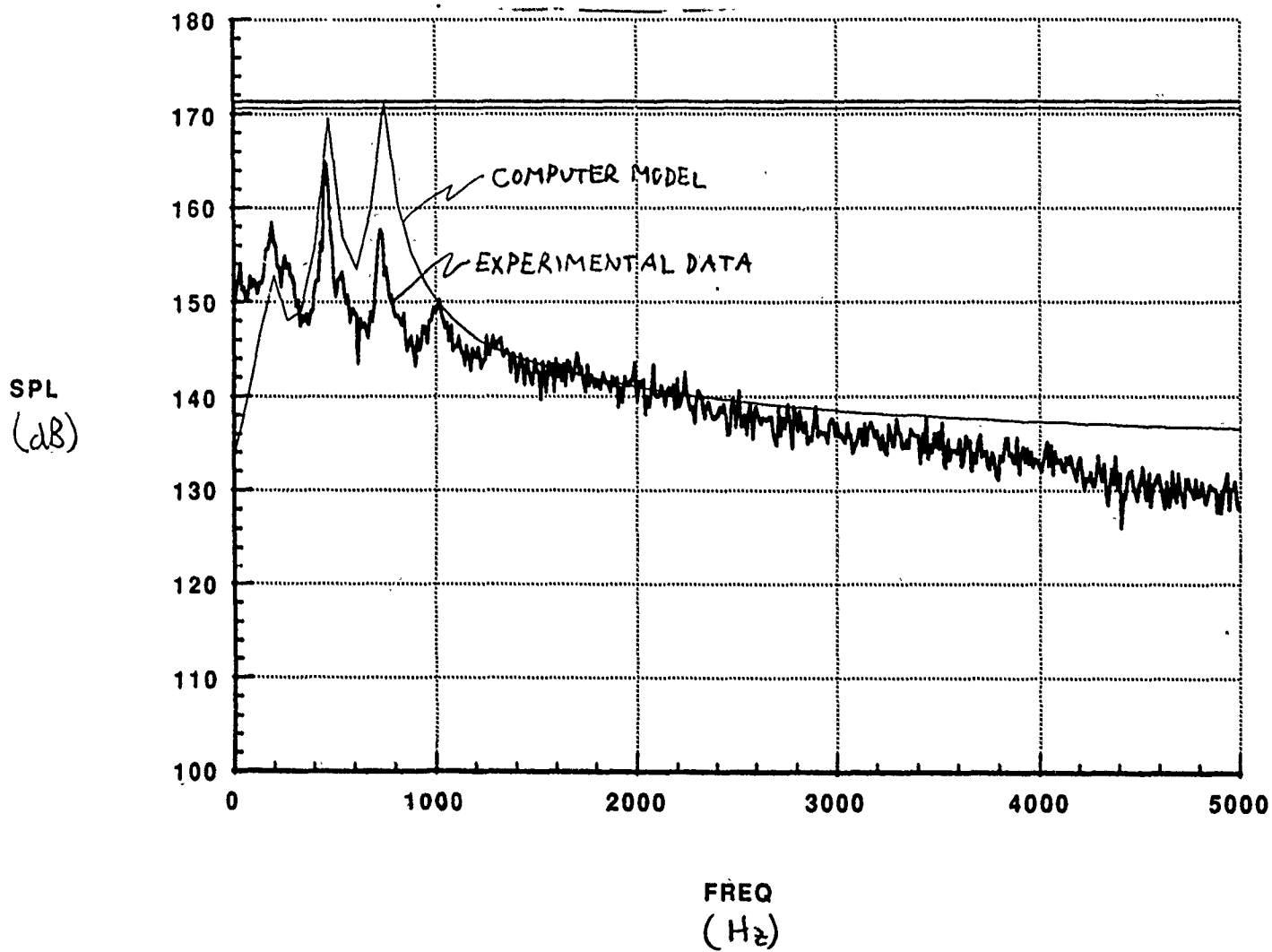


Figure 2. Mach No.=0.95

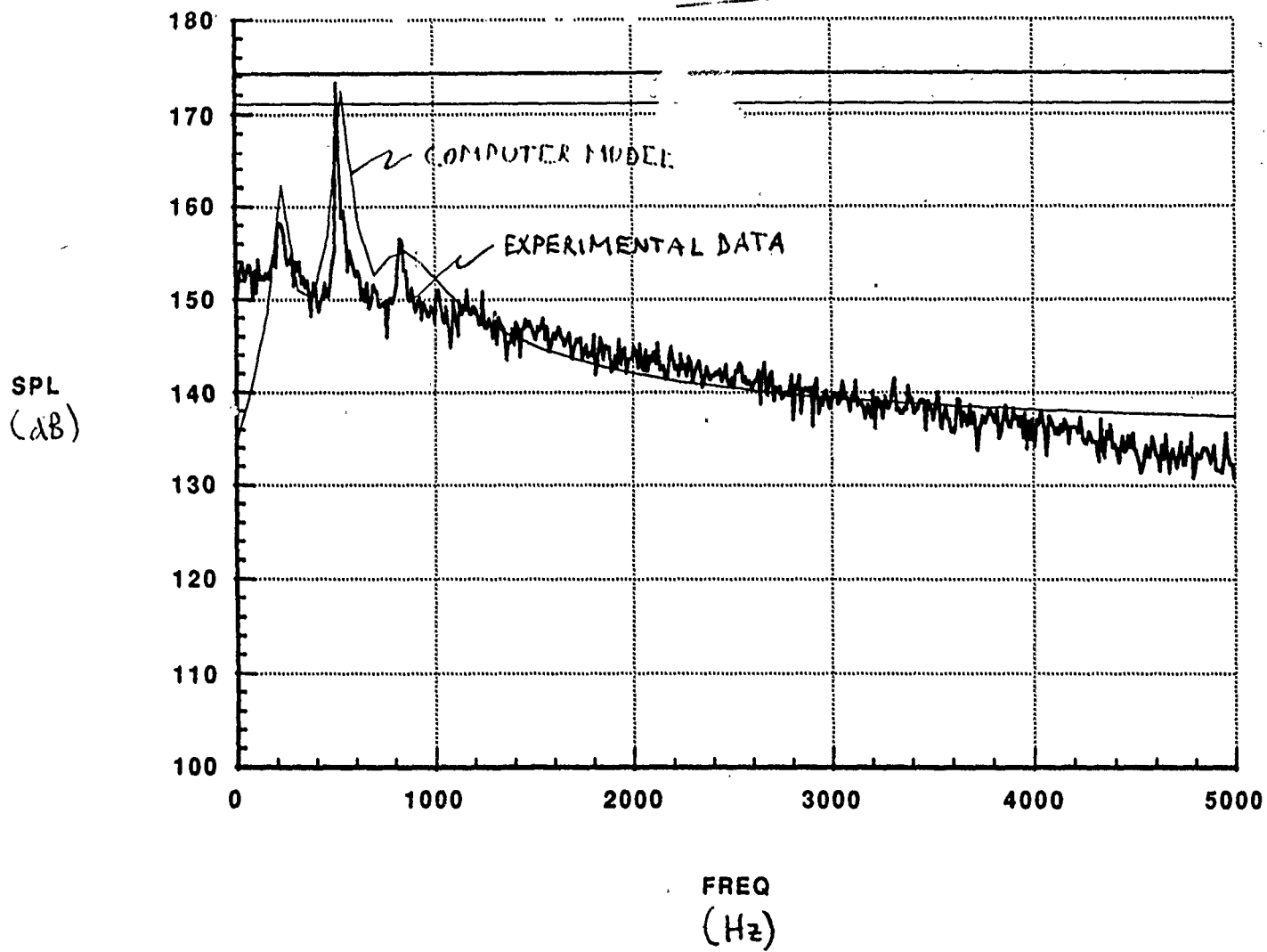


Figure 3. Mach No.=1.2

LIST OF REFERENCES

1. Chen, C.H. "Study of Subsonic and Transonic Flow Separation - With and Without Upstream Disturbances." Ph.D. Dissertation, The University of Tennessee, Knoxville, Tennessee, 1975.
2. Franke, M.E. and Carr, D.L. "Effect of Geometry on Open Cavity Flow-Induced Pressure fluctuations." AIAA Paper 75-492, 1975.
3. Heller, H.H. and Bliss, D. "The Physical Mechanism of Flow-Induced Pressure Fluctuations in Cavities and Concepts for their Suppression." AIAA Paper 75-491, 1975.
4. Vakili A.D. et.al. "Shear Flow Control Applied to Suppress Cavity Oscillations and Improve Store Separation." Weapons Carriage and Separation Workshop. Reno, Nevada, 5-6 April 1988.
5. Taylor, M. "Cavity Flow Oscillation Analysis and Water Table Simulation." M.Sc. Thesis, The University of Tennessee, Knoxville, Tennessee, 1989.
6. Karamcheti, K. "Acoustic Radiation from Two-Dimensional Rectangular Cut-outs in Aerodynamic Surfaces." NACA TN 3487, August 1955.
7. Roshko, A. "Some Measurements of Flow in a Rectangular Cutout." NACA TN 3488, August 1955.
8. Rossiter, J.E. "Wind Tunnel Experiments of the Flow Over Rectangular Cavities at Subsonic and Transonic Speeds." A.R.C. 3438.
9. Tam, C.R.W. and Block, P.T.W. "On the Tones and Pressure Oscillations Induced by Flow Over Rectangular Cavities." Journal of Fluid Mechanics. Vol. 89, Part 2, 1978, pp 1373-1399.

1990 USAF-RDL SUMMER FACULTY RESEARCH PROGRAM/
GRADUATE STUDENT RESEARCH PROGRAM

Sponsored by the
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by the
Research & Development Laboratories

FINAL REPORT

Monitoring an Axial Flow Compressor
for Rotating Stall and Surge

| | |
|--------------------|---|
| Prepared by: | John M. Sebghati |
| Academic Rank: | Master's Degree candidate |
| Department and | Aerospace Engineering |
| University: | University of Tennessee Space Institute |
| Research Location: | Arnold Engineering Development Center Arnold AFB, TN 37389 |
| USAF Researcher: | Jim Shelton |
| Date: | 9 September 1991 |

Monitoring an Axial Flow Compressor
for Rotating Stall and Surge

by

John M. Sebghati

ABSTRACT

Rotating stall and surge in an axial flow compressor are unsteady phenomena, which if left unattended could prove disastrous for the compressor and possibly the entire wind tunnel. Detecting the onset of these phenomena can, at times, prove to be difficult and requires constant monitoring.

Therefore, a study of the methods for monitoring the axial flow compressors in the 16 ft. transonic and the 16 ft. supersonic wind tunnels located at the Arnold Engineering Development Center (AEDC), Arnold Air Force Base was conducted for the detection of these unsteady phenomena. A method to complement or replace the existing use of strain gages mounted on a select number of rotor blades in each stage of the axial compressor was desired because of its slow response time. A variety of techniques were found, these include hot-wire, thermocouples, pressure transducers, acoustic sensors, and optical sensors. The use of high response static pressure transducers was found to be the most promising method but in light that it is still a new technology, this is not the time for its full implementation. A neural network based expert system tied into the current system would, for the time being, be the best possible solution and would provide a method for distinguishing between rotating stall, surge, and any other abnormalities, at a variety of operating conditions.

Acknowledgements

I wish to thank the Air Force Systems Command and the Air Force Office of Scientific Research for sponsorship of this research. Also, thanks should go to Research & Development Laboratories for aiding me in all administrative and directional aspects of this program.

The experience I received through this program has been of great importance to me, and thanks should be given where it is due. Dr. Ching F. Lo from the University of Tennessee Space Institute (UTSI) has provided me with the overall guidance and direction. Jim Shelton needs to be thanked for providing me with support and encouragement. Richard Womack and Troy Bisby must also be thanked for their help in some of the detailed aspects of my research. And finally, thanks must go to Brenda Warren for her help in tracking down much of the material that I needed.

I. INTRODUCTION:

Peak performance and efficiency have become very important in the use of axial compressors. The detection of unsteady phenomena like rotating stall and surge in axial flow compressors are of a major importance in achieving this. The speed of a compressor can vary over a wide range and at each speed there is a possibility of a surge or stall occurring, depending on the flow conditions present. Operating under these conditions has the potential of causing major damage to the axial compressor and possibly the entire tunnel.

The current method in use at the Arnold Engineering Development Center (AEDC) involves strain gages located on a number of blades in each compressor stage, and their constant monitoring by human operators. While a surge is easily detected using these strain gages, the rotating stall is somewhat more difficult to catch, especially by an operator without a lot of experience in this area. In order to make sure that a rotating stall is occurring, the operators typically wait several minutes to see if the abnormality picked up by the strain gages continues. On average it takes 15 - 20 minutes for the people watching over the compressors to notice and determine without any doubt that a rotating stall is taking place (Womack, 1991). On a few occasions a rotating stall has gone unnoticed for an extensive period of time on account of the change in the stress trace being subtle because the stall region was small. A method of detecting surge and rotating stall nearer its inception is desired in order to curtail this major cause of disturbances during normal tunnel operations.

AEDC has a large number of wind tunnels. The two largest, shown in Figure 1, are the 16 ft. transonic (16T) and 16 ft. supersonic (16S) tunnels which are the focus of this study. 16T is capable of providing a Mach range from 0.60 to 1.60 (Joyce, 1975), 16S is capable of generating Mach numbers from 1.60 to 4.00. They are closed-return, variable density and temperature wind tunnels that provide a 16 ft. wide by 16 ft. high by 40 ft. long test section. A diagram of the 16T axial flow compressor, C1, which is approximately 32 ft. in diameter can be found in Figure 2. It can provide a pressure ratio from a minimum of 1.11 to a maximum of 1.60. The four compressors, C2 - C5, of 16S are shown in Figure 3, the largest of which is 32 ft. in diameter. It is capable of generating pressure ratios from 1.20 to 7.58.

II. OBJECTIVES OF THE RESEARCH EFFORT:

A complete understanding of the unsteady phenomena of rotating stall and surge in an axial flow compressor is needed in order to make a decision on how to better detect the onset of these phenomena. There is a desire to implement a more accurate and quicker method of detection than the one currently being used in the axial flow compressors of the two large wind tunnels at AEDC, the 16 ft. transonic tunnel and the 16 ft. supersonic tunnel. It was my task as a participant in the 1991 Graduate Student Research Program (GSRP) to gather the information needed to form this decision and to make a recommendation as to what method should be used for the detection of rotating stall and surge in these two axial compressors.

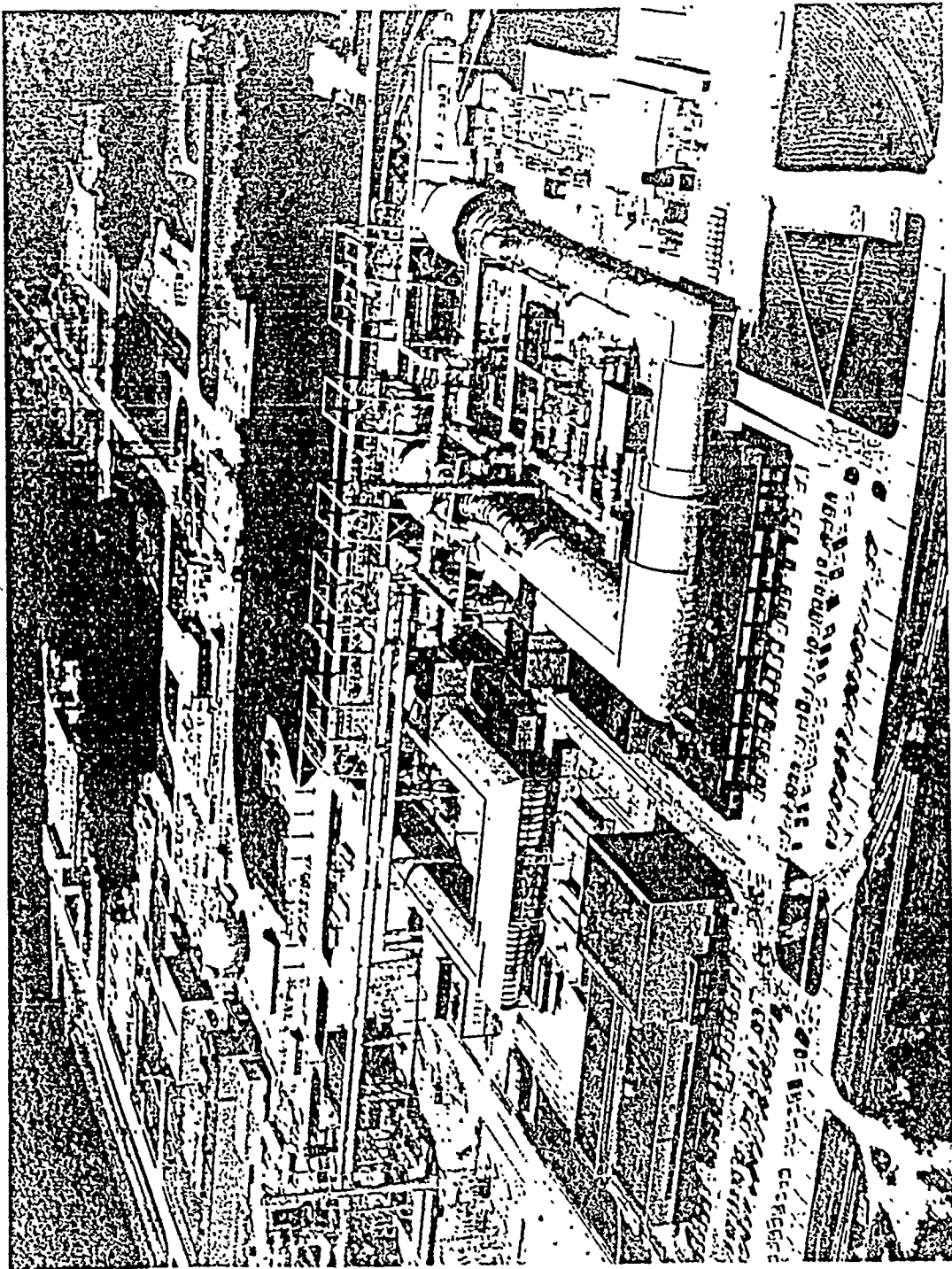
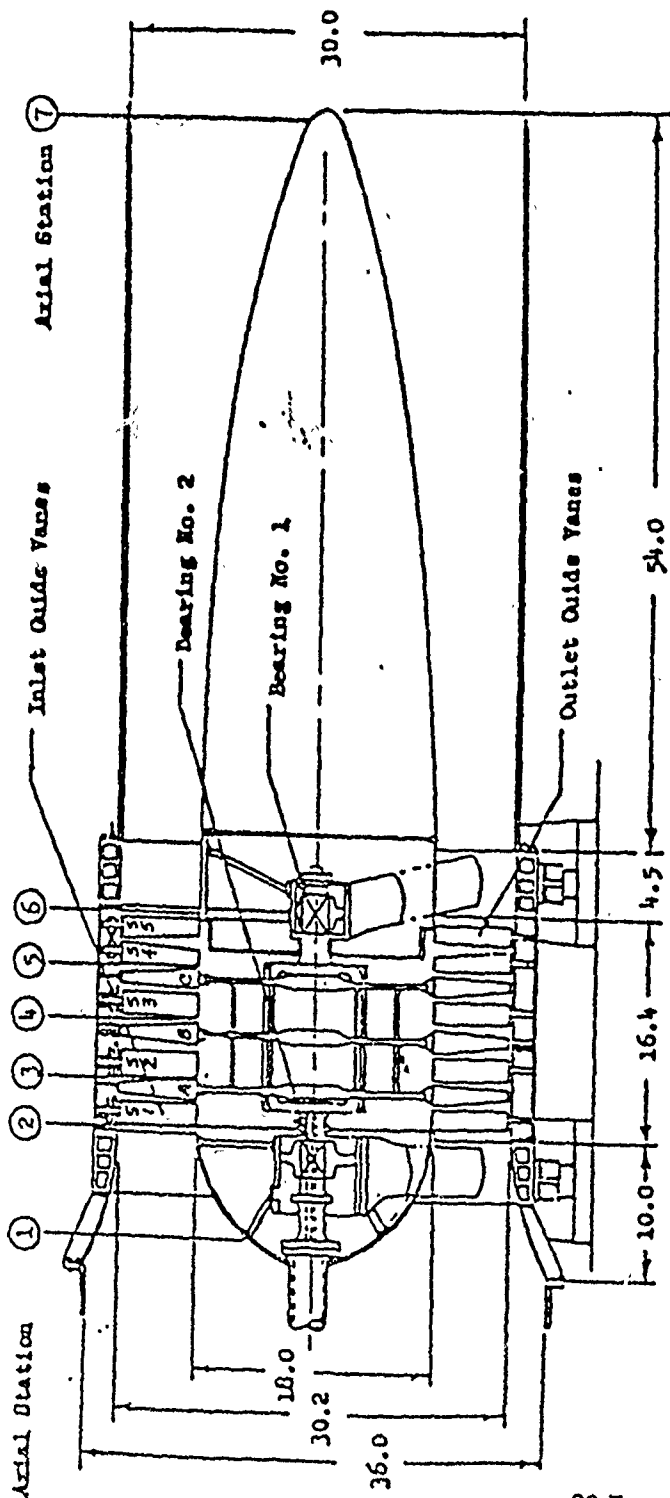


Figure 1. Propulsion wind tunnels.



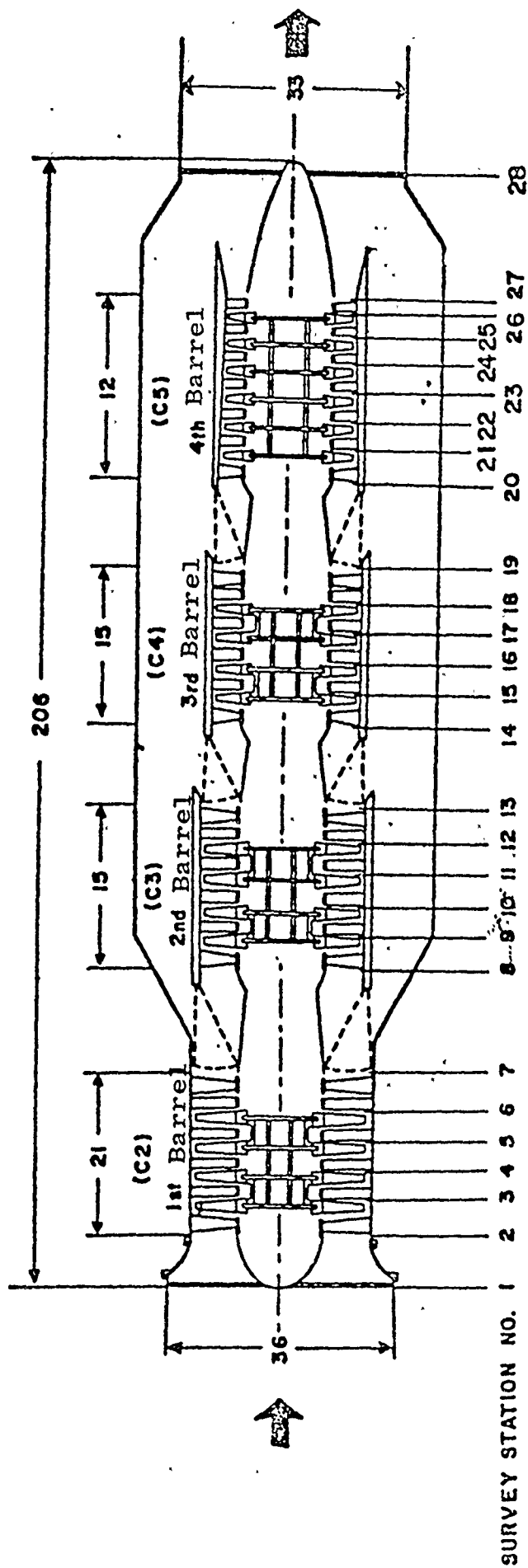
All Dimensions in Feet

Design Conditions

Tip Speed 542 ft/sec (Absolute)
 Hub Speed 565 ft/sec (Absolute)
 Design Pressure Ratio 1.385
 Design Inlet Volume Flow 200,000 cfs
 Flow Coefficient 0.47
 Work Coefficient 0.308
 Inlet Axial Velocity 442 ft/sec

Figure 2. Compressor C1

| Barrel | Hub Diameter | Tip Diameter | No. of Stages of Compression | Hub Speed Ft./Sec. | Tip Speed Ft./Sec. |
|--------|--------------|--------------|------------------------------|--------------------|--------------------|
| C 2 | 18 | 27.8 | 4 | 565 | 872 |
| C 3 | 18 | 25.4 | 4 | 565 | 797 |
| C 4 | 18 | 23.9 | 4 | 565 | 750 |
| C 5 | 18 | 22.3 → 20.9 | 6 | 565 | 700 → 656 |



All Dimensions in Feet
Not to Scale

a. Tunnel 16S Compressor (Four Barrel)

FIGURE 3. Compressor Dimensions and Longitudinal Survey Stations

III. RESULTS:

a. An extensive literature search was conducted at the AEDC library. Keywords like axial flow compressor, rotating stall, and surge were used in a computer-assisted literature search. Numerous articles, books, and on-base technical reports were found through this search. Two interviews were conducted with Dick Womack, an engineer at AEDC with many years of experience with the 16 ft. transonic and supersonic tunnels. Through this process a thorough understanding of the axial flow compressor, the unsteady phenomena of rotating stall and surge, and the two 16 ft. wind tunnels at AEDC was gained.

b. The purpose of an axial flow compressor is to increase the total pressure of the air through the use of shaft work. In order to do this in a highly efficient manner the compressor has been designed to operate near the surge point. With the introduction of some form of instability that changes the steady axisymmetric flow, like a fluctuation in pressure ratio, mass flow, rotational speed, or the turbulent boundary layer on the compressor blades or annulus, the possibility of an unsteady phenomena occurring is present (Greitzer, 1980).

If the compressor operating point moves beyond the surge point the flow is no longer steady. 16T primarily sees rotating stall and 16S surge. It has been reported that rotating stall always precedes a surge but that has not been seen to occur in 16S. If it does occur its duration would be on the order of milliseconds. A surge is defined as mass flow oscillating back and forth through the axial compressor. A rotating stall, on the other hand, consists of the flow across a number compressor blades becoming

detached due to an increased angle-of-attack from an instability upstream. This region of stalled flow rotates around the compressor usually between twenty and seventy percent of the rotor speed.

In order to detect rotating stall and surge one must look at what occurs during these two phenomena. During rotating stall compressor blades pass through the stalled region causing a large vibratory stress on the blades. In the stalled area there is a reduction in mass flow, change in pressure and temperature, and a possibility of some kind of acoustic resonance. A surge sees a complete reversal of mass flow which causes a high blade stress, a change in the pressure ratio, a noticeable temperature change in some compressors, and depending on the compressor an acoustic resonance may also occur. The magnitude of these depend greatly on the design of the compressor and a few may not be noticeable at all.

Methods of detecting rotating stall and surge use these changes. These methods include hot-wire, thermocouple, pressure transducer, strain gage, optical sensor, and acoustical sensor. The hot-wire can detect changes in mass flow (Laurence, 1952). They would be very good to use but are extremely fragile and expensive. Thermocouples are very effective and are currently used in many axial flow compressors, mainly small ones since they have a relatively constant temperature distribution. For it to detect temperature changes in large compressors a number of them would have to be located around the circumference and they would have to be extremely sensitive. This method has no guarantee of detecting a change right away in such a large compressor. A recent discovery is that the inception of a rotating stall can be detected in low and high speed axial compressors using high response static pressure transducers located on the casing wall about the circumference (Garnier et al, 1991; Inoue et al, 1991) . In order to have the earliest detection possible, these pressure transducers would

need to be located at each compressor stage. The only disadvantage with this method is that it is still in the research stage and needs time for further development.

Strain gages are a very good method, they detect the stress in the compressor blades (Fowler, 1979). A surge is easily detected because its stress trace is distinctive and it occurs suddenly. A rotating stall, though, is a bit more subtle. Strain gages also pick up normal operational changes which could be mistaken for a rotating stall. This is an effective method but not a quick one and the cause of the stress could be temporary and come from normal operations. The purpose of an optical sensor is to pick up blade vibrations which are caused by the passage of compressor blades through a region of stall and by flow reversal (Fowler, 1979). This method has the capability of replacing the strain gages but its downfall is the same in that the cause of the vibration would be unknown and thus is not a quick response. An acoustical sensor picks up the acoustic resonance of the blade vibration, therefore giving it the same disadvantages as the optical sensor (Parker, 1984).

There is a need for something to assist the compressor monitors in the discovery of rotating stall and surge in the compressors. Using an on-line artificial intelligence system would allow for quick detection of an abnormality and an analysis of the problem. If the system knew what the operating conditions would be during the test run, it would be able to differentiate between normal changes and abnormal changes. The use of a neural network based expert system hooked up to the sensors in the axial compressor would be able to weight all of the incoming data and use it in its final decision which would be presented to the operators (Liu, 1991; Lo et al, 1991). The operators could then review the reasons of the system for declaring an alarm and follow up on it if they agree.

IV. RECOMMENDATIONS:

a. The incorporation of a neural network based expert system into the monitoring process is currently the only thing that will speed up the detection of rotating stall and surge. Connecting the strain gages into the neural network based expert system would allow for a quicker detection of many of the problems that can occur to the axial compressor and it would speed up the correction of these problems. This system would also give a certain flexibility to the compressor operators in allowing them the further opportunity to analysis any unusual circumstances that may happen in the operation of the compressors.

b. The method involving the use of high response static pressure transducers is very promising and should be closely watched. If further research continues to prove that it can detect rotating stall in its early stages then it is just what AEDC needs for 16T and 16S. One question that still needs to be answered, though, is will it work in axial flow compressors the size of C1 - C5.

REFERENCES

Fowler, R.B., "Online Monitoring of Compressor Blade Stress: An Improved System," Arnold Engineering Development Center TR-79-7, Arnold Air Force Base, Tennessee, April, 1979.

Garnier, V.H., Epstein, A.H., Greitzer, E.M., 1991, "Rotating Waves as a Stall Inception Indication in Axial Compressors," ASME Journal of Turbomachinery, Vol. 113, pp. 290-302.

Greitzer, E.M., 1980, "REVIEW - Axial Compressor Stall Phenomena," ASME Journal of Fluids Engineering, Vol. 102, pp. 134-151.

Inoue, M., Kuroumaru, M., Iwamoto, T., Ando, Y., 1991, "Detection of a Rotating Stall Precursor in Isolated Axial Flow Compressor Rotors," ASME Journal of Turbomachinery, Vol. 113, pp. 281-289.

Joyce, C.R., "A Real-Time Monitor and Display System for the Propulsion Wind Tunnel Compressors" Master's Thesis, University of Tennessee, Knoxville, Tennessee, 1975.

Laurence, J.C., Landes, L.G., "Auxiliary Equipment and Techniques for Adapting the Constant Temperature Hot-Wire Anemometer to Specific Problems in Air-Flow Measurements," NACA TN 2843, Nov. 1952.

Liu, W., "An Expert System for Predicting Flutter in Turbomachines," Presented at the AIAA/SAE/ASME/ASEE 27th Joint Propulsion Conference, Sacramento, CA, June 24-26, 1991.

Lo, C.F., Shi, G.Z., "Neural Network Based Expert System for Compressor Stall Monitoring," Presented at the AIAA/SAE/ASME 27th Joint Propulsion Conference, Sacramento, CA, June 24-26, 1991.

Parker, R., 1984, "Acoustic Resonances and Blade Vibrations in Axial Flow Compressors," Journal of Sound and Vibration, Vol. 92, pp. 529-539.

Womack, R., Private Correspondence, Arnold Engineering Development Center, Arnold Air Force Base, Tennessee, 14 June 1991 and 15 July 1991.

DEMONSTRATION OF A NEW FINITE-RATE CHEMISTRY CFD CODE

Paul Vitt
Graduate Student Research Assistant
Arnold Engineering Development Center

Abstract

This paper presents the procedure and test problem results for the initial demonstration of the GASP computational fluid dynamics code at AEDC. The code was converted to run on a Silicon Graphics IRIS workstation, and solve test cases and current problems of interest. These problems, a high back pressure nozzle and a high area ratio nozzle, were used to demonstrate the abilities of GASP. Some difficulties were encountered with convergence and small upper CFL number limits.

Nomenclature

| | |
|-------------------|--|
| CFL | Courant, Freidrichs, Lewy Number |
| e | energy |
| F | general function vector |
| F, F _v | inviscid, viscous flux vectors |
| h | enthalpy |
| I | identity matrix |
| J | transformation Jacobian |
| M | Mach number |
| P | pressure |
| T | temperature |
| u | velocity vector |
| V | species diffusion velocity |
| W, ω | vector, species source terms |
| x | Cartesian coordinate vector ($x_{1,2,3}$) |
| λ | eigenvalues |
| ρ, ρ | total, species densities |
| τ | viscous stress tensor |
| ξ | characteristic coordinate vector ($\xi_{1,2,3}$) |

1. Introduction

Methods for numerically solving compressible flows are improving rapidly, meaning

that larger and more complex problems can be solved with them. Since computational fluid dynamics (CFD) plays such a large role in modern testing and design, it is necessary for facilities such as Arnold Engineering Development Center (AEDC) to be proficient in this technology area. This research project involved setting up and demonstrating some of the capabilities of one such new code, GASP (developed by R. Walters, Virginia Polytechnical Institute). At the same time, this project was to be used as an introduction for the author to the methods and technology of modern computational fluid dynamics.

GASP was developed on Cray computers, and although generally compatible with other systems, some modification was required. Hence, the first step in this study was to modify the code to run on the target system, a Silicon Graphics IRIS system. After these modifications were made, a number of test cases which are provided with the code were run, to ensure that the system was properly configured. The final step in this preliminary demonstration was to input several simple problems that are of interest to the Propulsion Technology group at AEDC.

The conversion to the IRIS system was successfully accomplished. All test cases ran using the modified code except for two, both of which had Cray binary grids which were untranslatable by the IRIS-C language. In addition, there were three test problems that were of interest to the Propulsion technology group: a high back pressure rocket nozzle, a very high area ratio (1030:1) rocket nozzle, and a backward facing step. Both nozzle problems achieved suitable demonstration results (i.e., the code could handle the problems). The backward facing step caused difficulties that were not overcome during the course of this research due to limited time. GASP is a code that has many options for solution methodology. An improper choice of these may have led to the difficulties. A complete examination of these options may resolve the problem.

2. Computational Method

2.1 Mathematical Background

The new code that was to be demonstrated is called GASP: General Aerodynamic Simulation Program. The code is an upwind finite volume solver for the Navier-Stokes equations:

$$\frac{\partial}{\partial t} \left(\frac{Q}{J} \right) + \frac{\partial (\tilde{F} - \tilde{F}_v)}{\partial \xi} = \frac{W}{J} \quad (1)$$

where the conserved variable and source vectors are given by:

$$Q = \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_N \\ \rho u_{\xi_1} \\ \rho u_{\xi_2} \\ \rho u_{\xi_3} \\ \rho e \end{bmatrix}, \quad W = \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_N \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

and the flux vectors by:

$$\tilde{F} = \frac{\nabla \xi}{J} \begin{bmatrix} \rho_1 \tilde{u}_\xi \\ \vdots \\ \rho_N \tilde{u}_\xi \\ \rho \tilde{u}_\xi u_{\xi_1} + \xi_{x_1} P \\ \rho \tilde{u}_\xi u_{\xi_2} + \xi_{x_2} P \\ \rho \tilde{u}_\xi u_{\xi_3} + \xi_{x_3} P \\ \rho h \tilde{u}_\xi \end{bmatrix}, \quad \tilde{F}_v = \frac{\nabla \xi}{J} \begin{bmatrix} -\rho_1 V_1 \\ \vdots \\ -\rho_N V_N \\ \tau_{\xi_1} \\ \tau_{\xi_2} \\ \tau_{\xi_3} \\ \sum_{j=1}^N \rho_j h_j V_j \\ k (\nabla \xi \nabla T) \end{bmatrix} \quad (3)$$

The terms in the vectors are noted in the nomenclature section, and vibrational nonequilibrium has been neglected here (but not in GASP). GASP has a number of methods with which this system of equations may be solved. Algebraic models are used for the chemical source terms, viscosity effects (Blottner curve fits; Sutherland's Law), thermal conductivity and laminar mass diffusion. Turbulence effects are modelled by either the

Baldwin-Lomax algebraic model (which was tested successfully, although results are not included in this report) or with one of several k-epsilon methods. It is desirable to include k-epsilon turbulence studies in the code demonstration, but time restrictions for this study prevented accomplishing. The code is spatially discretized using upwind finite volume techniques (central differencing is also an option), which uses the integral formulation of equation (1) to solve over a volume enclosed by a boundary Ω through which the fluxes are evaluated. The fluxes can be treated in different ways. First, if equation (1) is written as:

$$\frac{\partial Q}{\partial t} + \frac{1}{J} \left[\frac{\partial F(Q)}{\partial \xi} - W \right] = 0 \quad (1a)$$

then it is solved as:

$$\frac{\partial Q}{\partial t} + \frac{1}{J} \left[\frac{\partial F}{\partial Q} \right] \frac{\partial Q}{\partial \xi} = \frac{W}{J} \quad \rightarrow \quad \frac{\partial Q}{\partial t} + \frac{A}{J} \frac{\partial Q}{\partial \xi} = \frac{W}{J} \quad (1b)$$

where the matrix $[A]$ has to be inverted (implicit solution) or solved explicitly. For time accurate solutions, a Jameson-style Runge-Kutta scheme can be specified, and this will solve for Q explicitly. For steady problems, approximate factorization or Lower-Upper decomposition solve for Q implicitly. Only implicit methods were used in this study. $[A]$ is the Jacobian matrix of the fluxes with respect to the conserved variables. In order to increase the efficiency of the solver, split-flux methods were developed. The split flux options are Steger-Warming, Van Leer ([2],[3]), or Roe's flux differencing methods ([4]). The split flux methods increase efficiency by forming the fluxes into separate waves that progress in opposite directions. The methods split $[A]$ into $A = A^+ + A^-$ and F into F^+ and F^- , where:

$$A^+ = A + \lambda I \quad , \quad A^- = A - \lambda I \quad (4)$$

Split flux methods are naturally dissipative, which adds numerical smoothing equivalent to

the explicitly added artificial dissipation in full flux (central difference) methods. This damps nonlinear instabilities in the method. Flux limiters are usually used to suppress dispersive "wiggles" around jump conditions, creating the ability to capture sharp shocks.

2.2 Program Modifications

There were few modifications that had to be made to GASP in order to implement it on a Silicon Graphics IRIS 4D-240 computer. The major reason that GASP was implemented on the IRIS system instead of the mainframe Cray at AEDC was that the Cray only has 1.8 million words of memory -- too limited to handle GASP problems. The IRIS system has 16 million words (64 megabytes of RAM) of memory, which in principle is large enough except for cases involving the larger chemistry models. The major modification came in changing several subroutines to replace the Cray's internal subroutines. The major problem encountered in transforming GASP to the IRIS is a lack of accuracy. A Cray runs with 8 byte accuracy, while the IRIS system has only 4 bytes. It is recommended (that if GASP is continued to be used on the IRIS) that the program be double precision. It was originally thought that this problem could be solved through the use of implicit statements, but it is possible that the functions (which are not currently double precision) caused some of the problems with the third test case, and with the CFL number in general. With modifications, most of the test cases ran as they should (some of the cases had Cray binary packed grids, which could not be used on the IRIS without modification, so these were not run). Most of the problems could not initially be run at the CFL number recommended for the Cray, although in some cases the CFL number could be increased as the solution converged. At that point in time, this was not considered a problem -- it only hindered accuracy slightly, and computational time was not much of a concern, with the IRIS being (for all practical purposes) a dedicated machine. It is now believed that one possible reason that the third

test problem did not run successfully was because of the accuracy problems in inverting the matrix on the IRIS. This also caused a significant increase in computational time for all of the test problems.

3. Test Problems

The Propulsion Technology section was interested in achieving solutions for three problems as noted in the introduction. The first involved a perfect gas high back pressure nozzle (ref. [5]). The geometry for this case is shown in figure 1, and the grid used is shown in figure 2, with boundary conditions in Table 1. The actual grid was a three-dimensional axisymmetric grid, as can be seen in figure 1. The interest in this problem stemmed from solutions from two other

Table 1: Boundary Conditions for Test Problem 1

| Boundary | Boundary Condition |
|------------------|--|
| Inflow Plane | Specified Conditions: P, T, ρ, u fixed from file |
| Outflow Plane | $P_{\text{back}} = 5.40\text{e}+04 \text{ Pa}$ |
| Walls | No-slip, adiabatic |
| Sideplanes | Tangency condition |
| CFL number range | 0.002 to 0.5 |

codes, PARC and RPLUS, which showed a recirculation zone that started inside the nozzle, and ran out along the centerline.

The second was a perfect gas high area ratio (1030:1) nozzle, which would be used for

low orbit applications (references [6],[7],[8]). The test problem geometry is shown in figure 3. One reason for the interest in the GASP solution was that GASP has the ability to solve the transonic throat region elliptically, and then turn and march using parabolized-Navier-Stokes (PNS) approximations to solve the supersonic nozzle portion. Grids were developed for this approach in two blocks: the first encompassed the combustion chamber outlet to just downstream of the throat, and a second grid discretizing the supersonic region. These are shown in figures 4 (entire system) and 5 (transonic and early supersonic region). The chosen boundary conditions are listed in Table 2.

Table 2: Boundary Conditions for Test Problem 2

| Boundary | Boundary Condition |
|------------------|--|
| Inflow Plane | $P_{\text{tot}} = 2.45 \times 10^6 \text{ Pa}$, $\rho = 1.0655$ $T_{\text{tot}} = 3400 \text{ K}$, $M = 0.14$ |
| Outflow Plane | Extrapolated |
| Walls | No-slip, adiabatic |
| Sideplanes | Tangency condition |
| CFL number range | 0.1 to 1.0 |

The final problem that was to be solved was subsonic flow over a backward step (ref. [9]). This problem has a large recirculation zone, and should provide a rigorous test of the turbulence models. It also would try out injection boundary conditions -- which, as it turned out, GASP handled without any apparent problems. For reference, the problem is shown in figures 6, 7a and 7b, which show the experimental setup, the basic case and also

the injection modelling case. The inflow and outflow boundary conditions were varied, but the solution always diverged before it could converge very far, and any convergence achieved was done with extremely small local CFL numbers ($<.01$).

4. Test Problem Results

4.1 High Back Pressure Nozzle

The results for the first test problem are shown in figures 8 and 9, for Mach number and velocity magnitude, respectively. These results are not converged solutions: they were only converged far enough to demonstrate that GASP would not recirculate, which the figures clearly show that it did not. The solution is second order, and was converged three orders of magnitude, which is good for an approximate engineering solution. For a complete solution, the downstream boundary (exit plane) should be moved to a point beyond where the flow becomes totally subsonic, in order to have a representative boundary condition, and not influence the solution as can be seen at the right hand side of figure 8. From figure 8 can also be seen the large shear layer (the problem was run viscous, but confined to laminar boundary layers), and the beginning of a correct shock "diamond" solution, which was captured using Roe's method of flux differencing.

4.2 High Area Ratio Nozzle

This problem turned out to be more extensively studied than the first problem. Three perfect gas solutions are presented in figures 10, 11, and 12 for Mach number and temperature: an inviscid Roe's method first-order PNS (fig. 10), a laminar Van Leer split flux method second-order PNS (fig. 11), and a laminar elliptic Van Leer solution (fig. 12). All of the solutions except the elliptical problem were converged four orders of magnitude. The elliptical solution is not truly an elliptical solution, but rather is a global solution of the supersonic flowfield. It was arrived at by using the converged second-order Van Leer

solution as an initial condition, and converged from that point. Time constraints meant that a fully global solution could not be calculated, so the results for the elliptic case are merely an improvement on the marching solution. The development of the boundary layer can be seen as the solution progresses from figure 10 to 12, with perhaps the most realistic solution (when compared to [6]) being the global improvement on the marching solution: fig. 12. The oscillations in the solutions near the centerline exit are due to the large cell size in this area. A direct comparison between the solutions is shown in both figures 13 and 14, for Mach number and pressure coefficient respectively (note: the throat of the nozzle is at an axial location of 0.0 m).

5. Conclusions

The object of this research was to enable GASP to run problems on an IRIS system, and to demonstrate the ability of the program to solve the test cases and also test problems of interest in the Technology group. The program successfully ran on the IRIS system, although the accuracy of the results was probably compromised. Solutions to most of the test cases and also to two out of three test problems were generated to acceptable solutions: in one test problem, GASP converged to a solution that had eluded two other upwind codes. GASP also successfully demonstrated the ability to globally calculate a solution in one zone and then turn and march using parabolized Navier-Stokes in a downstream supersonic zone.

To further test GASP on the IRIS, the recommendation that the intrinsic functions in the program be double precision must be made, in order to investigate the possibility that the convergence problems are machine accuracy related. As a final note, the last objective of this research was to act as an introduction for the author to CFD techniques, which it accomplished very successfully.

7. References

1. Walters, R. W., et al. "A Users Guide to GASP". NASA Langley Research Center, 1990.
2. Van Leer, B. and Turkel, E. "Flux-vector splitting and Runge-Kutta Methods for the Euler Equations". NASA Contractor report 172415, 1984.
3. Van Leer, B. et al. "A Comparison of Finite Volume Flux Vector Splitting for the Euler Equations". AIAA-85-0122, 1985.
4. Roe, P. "Characteristics-based Schemes for the Euler Equations". Ann. Rev. Fluid Mech. 18:337-65, 1986.
5. Bloomer, H., Antl, R., Renas, P. "Experimental Study of Effects of Geometric Variables on Performance of Contoured Rocket-Engine Exhaust Nozzles". NASA Technical Note D-1181, 1962.
6. Lankford, D. "Application of Conchas-Spray to Rocket Nozzle Analysis". AIAA-87-2127, 1987.
7. McAmis, R., Lankford, D., Torick, R. "Application of a Chemically Reacting Navier-Stokes Code for Predicting Nozzle Performance of a High Area Ratio Liquid-Propellant Rocket Engine". AIAA-91-0307, 1991.
8. Pavli, A., Kacynski, K., Smith, T. "Experimental Thrust Performance of a High-Area-Ratio Rocket Nozzle", NASA Technical Paper 2720, 1987.
9. Wu, M., Walterick, R., et al. "Turbulent Diffusion Flame Properties Behind a Step". AIAA-91-0079, 1991.

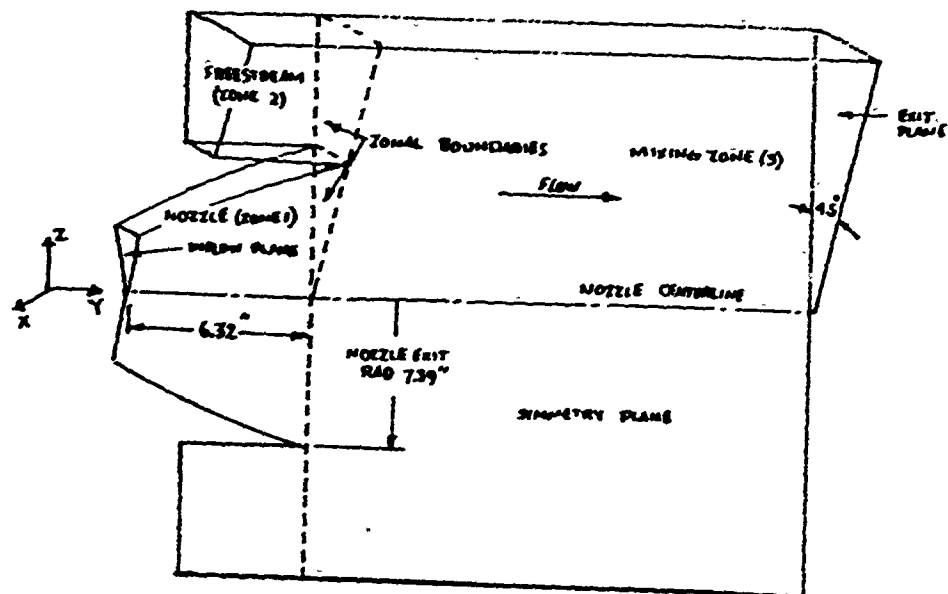


Figure 1: Test problem 1 geometry

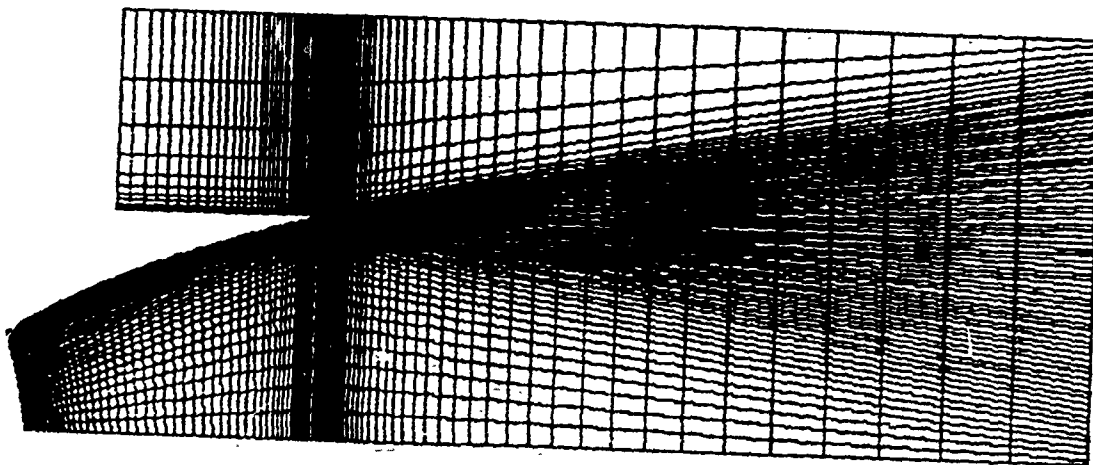


Figure 2: Test problem 1 grid layout

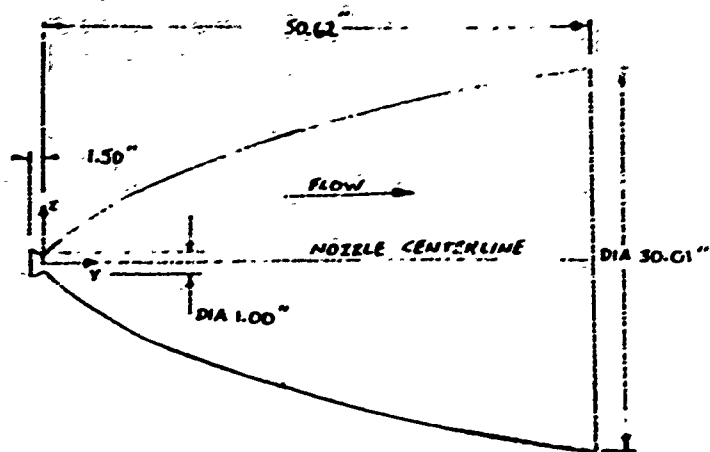


Figure 3: Test problem 2 geometry

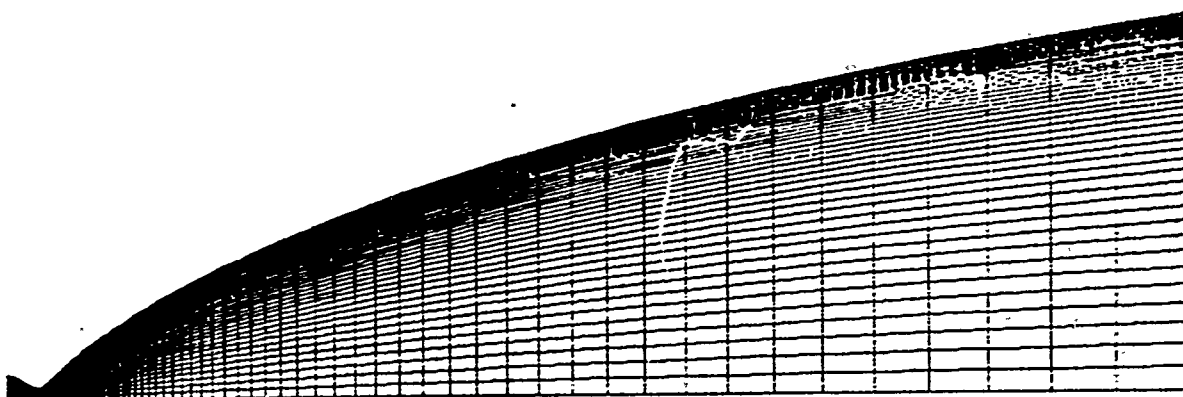


Figure 4: Test problem 2, entire problem grid layout

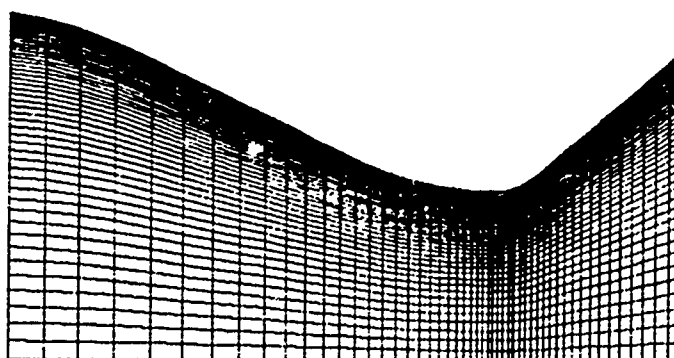


Figure 5: Test problem 2, transonic throat region grid detail

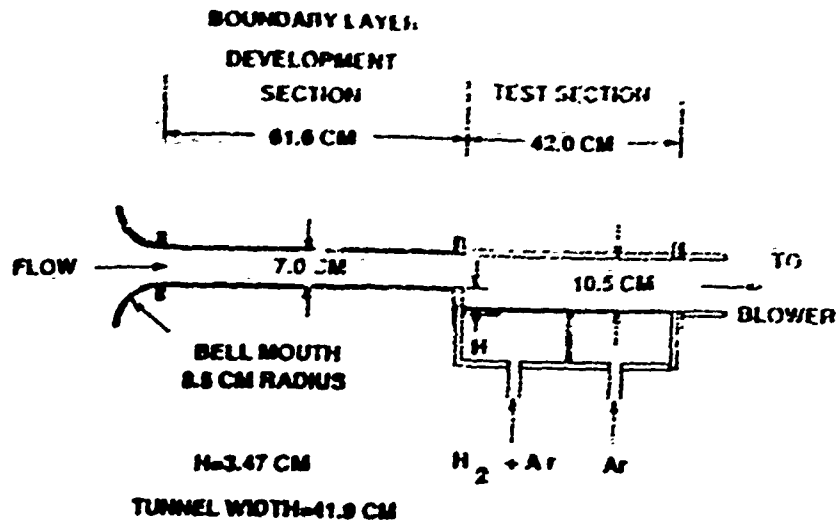


Figure 6: Test problem 3 experimental layout

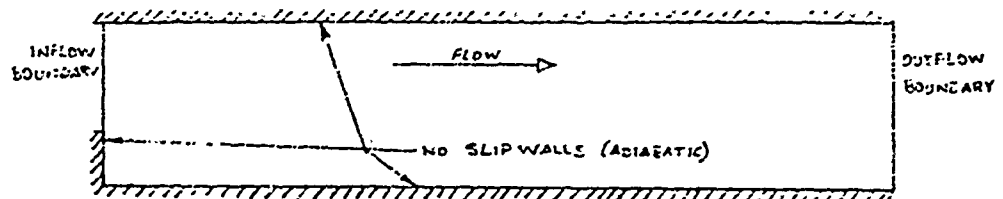


Figure 7a: Test problem 3 basic geometry layout

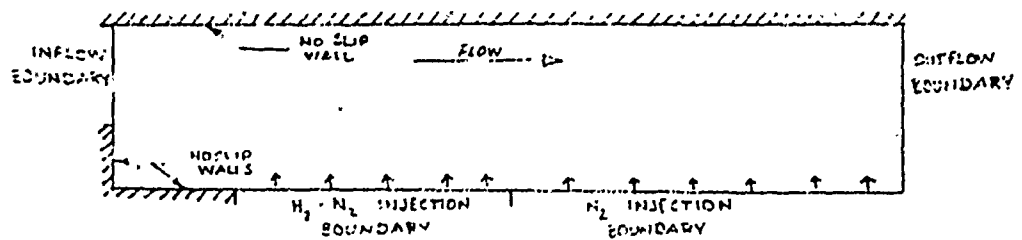


Figure 7b: Test problem 3 injection modelling geometry layout

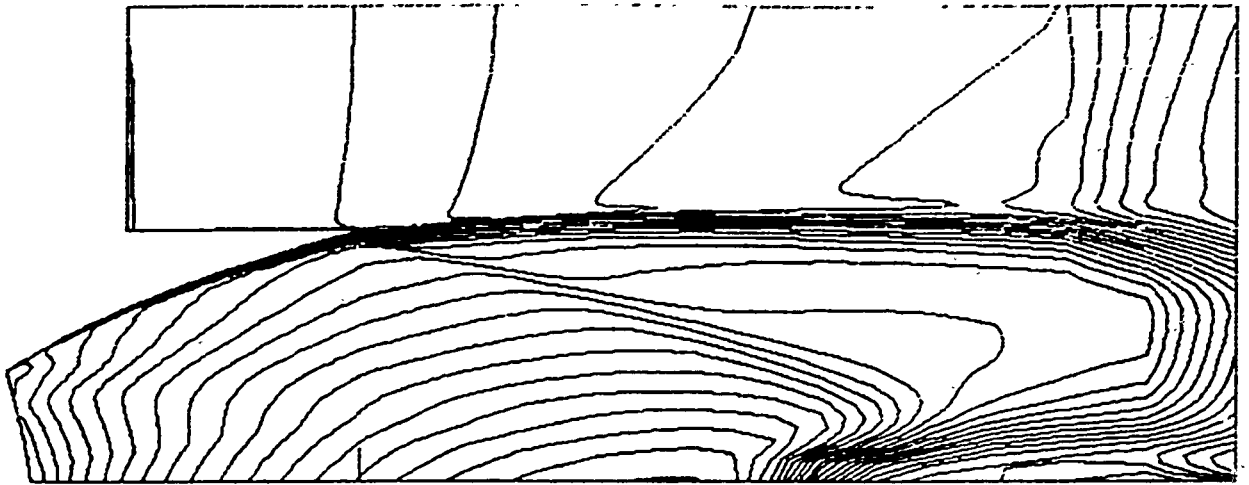


Figure 8: Test problem 1 Mach number contours

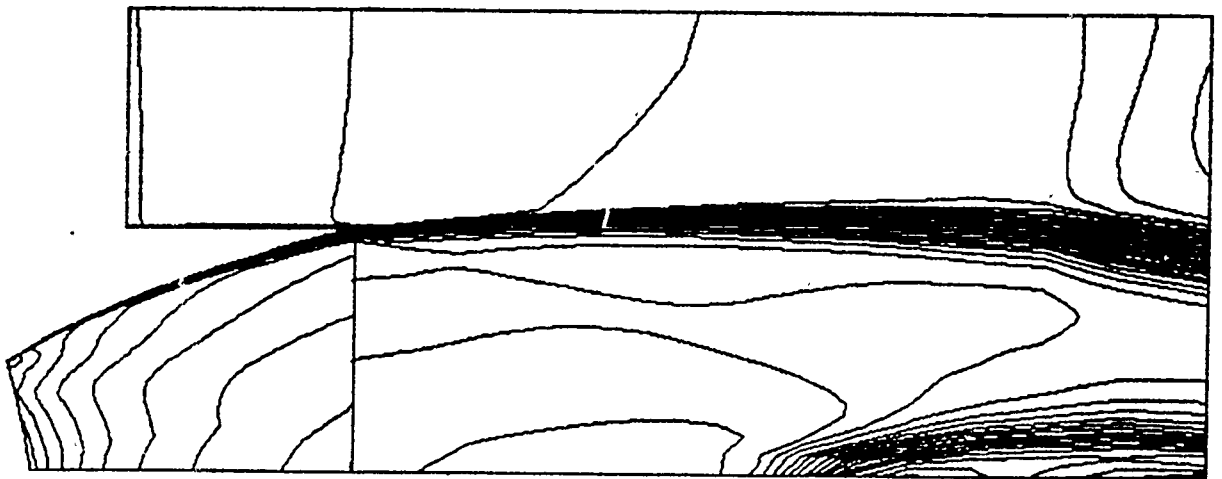


Figure 9: Test problem 1 velocity magnitude contours

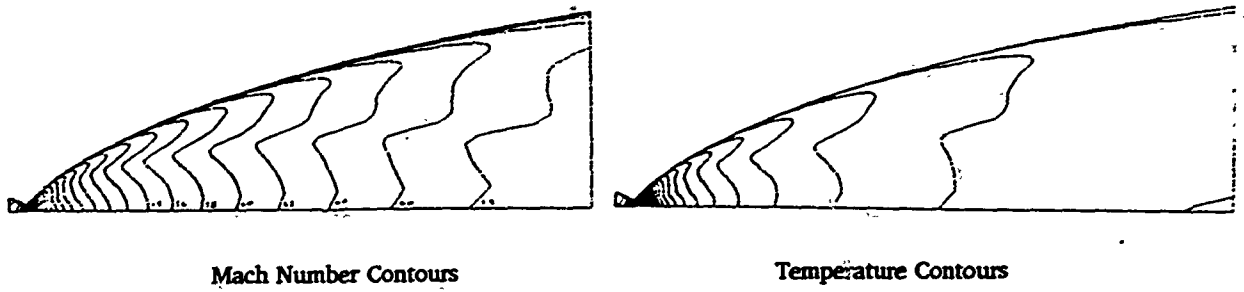


Figure 10: Test problem 2 first order Roe solutions

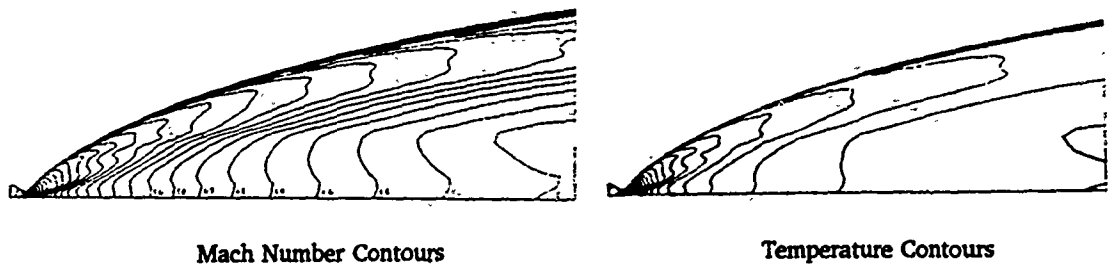


Figure 11: Test problem 2 second order Van Leer solutions

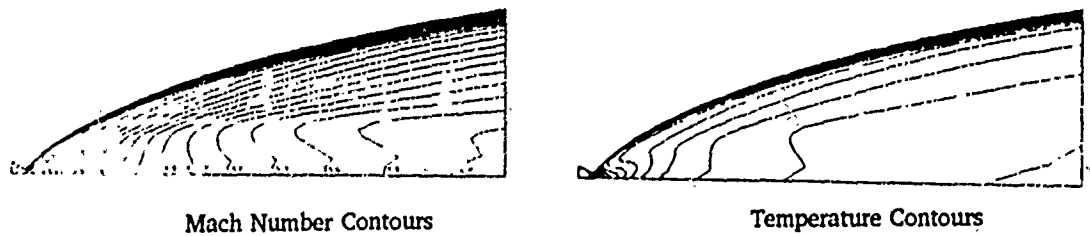


Figure 12: Test problem 2 global Van Leer solutions.

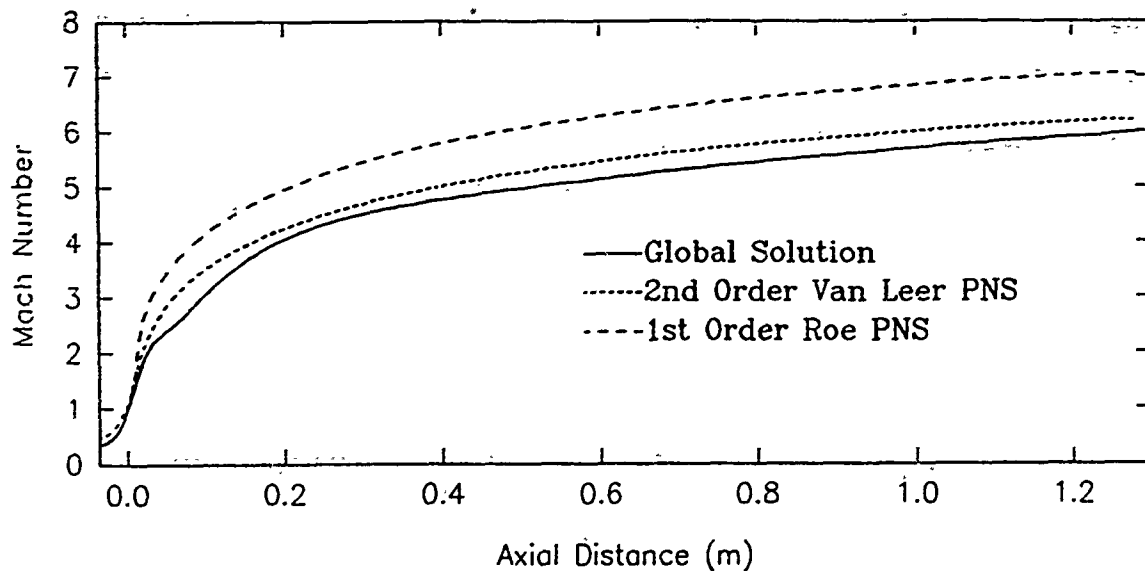


Figure 13: Test problem 2 solution comparison: Mach number

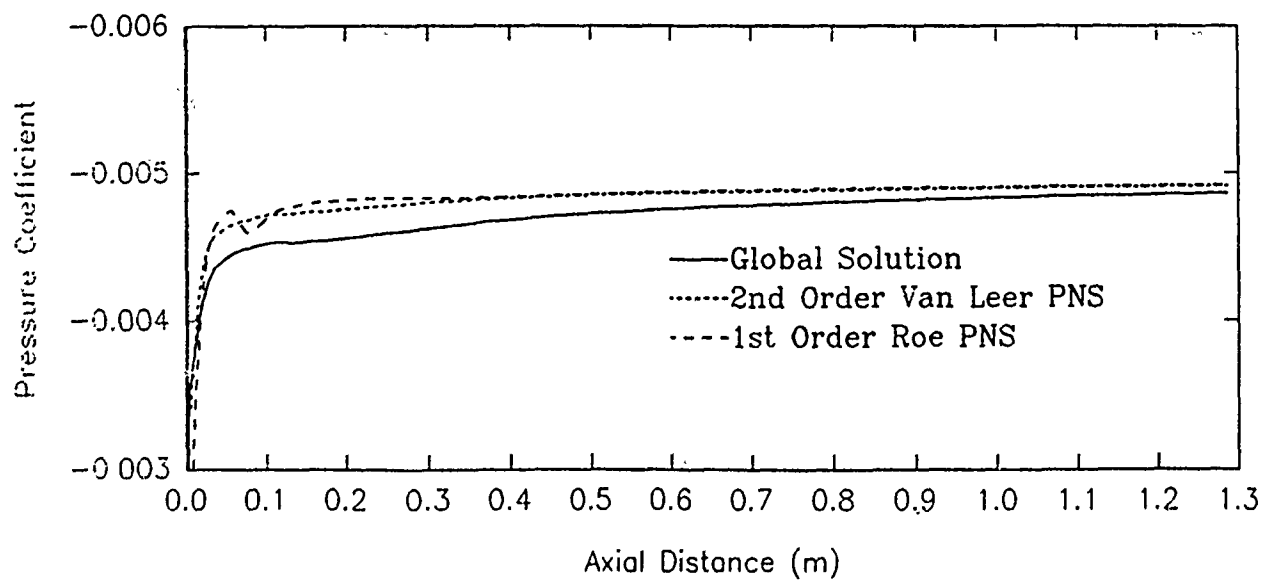


Figure 14: Test problem 2 solution comparison: pressure coefficient along the edge wall of the nozzle

**BOUNDARY ELEMENT ACOUSTIC ANALYSIS OF THE EXHAUST GAS
MANAGEMENT SYSTEM IN THE ARNOLD ENGINEERING DEVELOPMENT
CENTER AEROPROPULSION SYSTEMS TEST FACILITY**

Michael A. Weaver, MSAE
School of Aerospace Engineering
Georgia Institute of Technology

Abstract

The boundary element method was used to numerically determine the acoustic response of cylindrical models approximating the Exhaust Gas Management System in the Aeropropulsion Systems Test Facility at the Arnold Engineering Development Center. Longitudinal, tangential, and radial mode responses were calculated using two different model resolutions.

Introduction

This project was funded under the United States Air Force Office of Scientific Research 1991 Summer Research Program for Faculty and Graduate Students. The author is pursuing a PhD in computational fluid-structure interaction with the School of Aerospace Engineering at the Georgia Institute of Technology, where he has also obtained an MSAE degree. Work for this project was carried out at the Arnold Engineering Development Center (AEDC) in Tullahoma, Tennessee. The time frame for this project was the ten week period from 23 June 1991 to 31 August 1991.

The focus of this study is an acoustic interaction phenomenon which occurred at the Aeropropulsion Systems Test

Facility (ASTF), located at AEDC. A computational analysis was begun in support of work already conducted at AEDC on the problem. Guidance was provided by Dr. Ralph R. Jones from Sverdrup Technology Inc/AEDC Group (Sverdrup).

Background

ASTF provides the means to test full-scale aircraft propulsion systems under real-time environmental conditions simulating take-off, climb-to-altitude, multi-speed maneuvers, descent, and landing. The facility is managed and operated by Sverdrup.

The operation of the ASTF system (referring to Fig. 1) begins with external atmospheric air entering the compressors. Upon exiting the compressors, the air is heated or cooled as needed. Airflow is then guided into the test cell, where the test articles (typically an aircraft engine with its inlet) are placed. The airflow, including engine exhaust, then proceeds through an Exhaust Gas Management System (EGMS) appropriately chosen for the test article. From the EGMS, airflow enters the exhaust cooler, where it is cleaned and the temperature lowered for venting to the atmosphere.

The EGMS captures engine exhaust through an articulated front panel connected to a conical diffuser followed by a short circular cylinder. The articulated front panel allows the EGMS entrance to match various engine nozzle geometries. The configuration in Fig. 2 shows the EGMS for use with a 2-D, vectored thrust engine nozzle. In the side view shown,

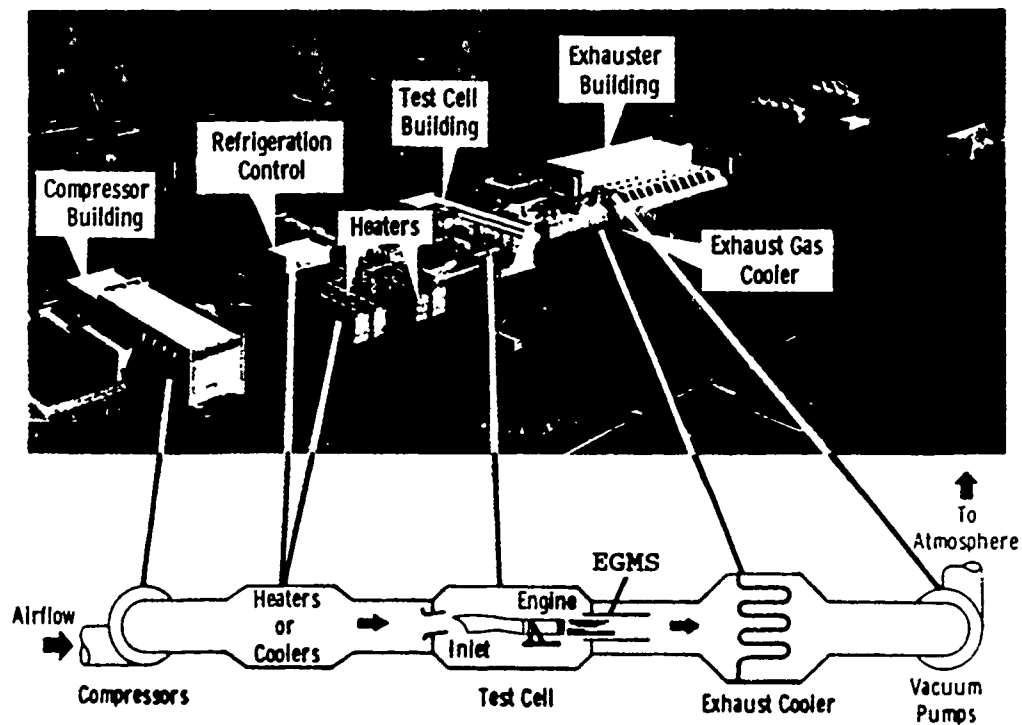


Fig. 1 - Operation of the Aeropropulsion Systems Test Facility

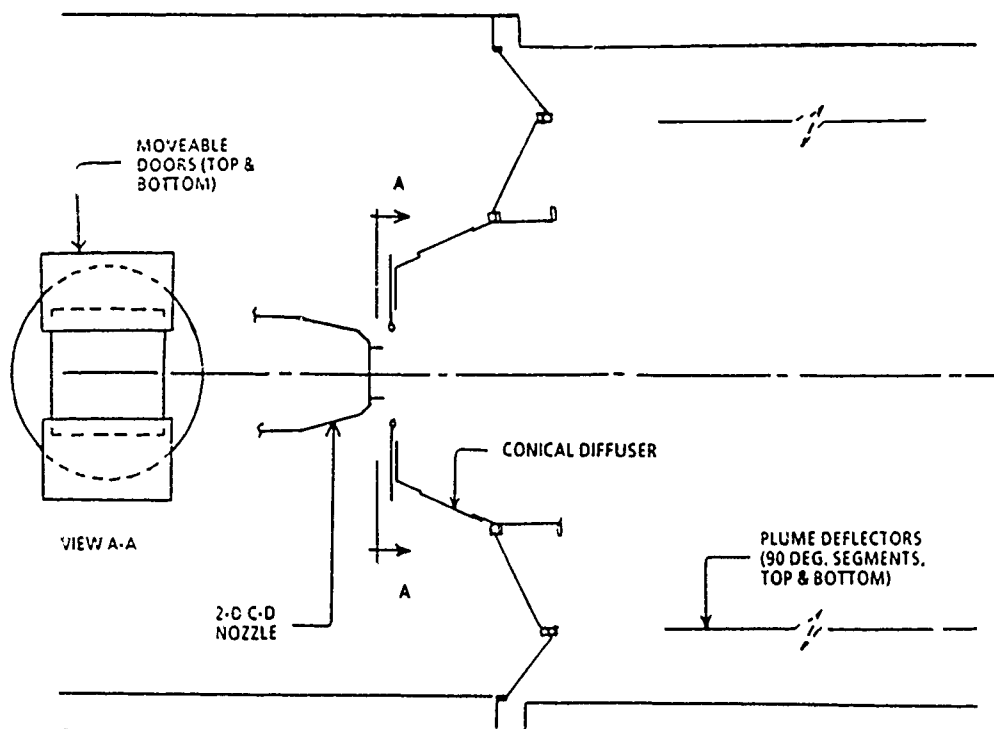


Fig. 2 - Exhaust Gas Management System for a 2-D, vectored thrust nozzle

the moveable doors traverse in the vertical direction, thereby increasing or decreasing the capture area and vertically translating the opening.

In 1989, during the testing of a 2-D, vectored thrust engine, anomalously high acoustic noise levels were produced in the EGMS. Two discrete tones were measured at 85 Hz and 135 Hz. Sound pressure levels were high enough (160 dB to 170 dB) that further testing could not proceed without imminent structural damage to the ASTF.

Attempting to alleviate this problem, several unsuccessful modifications were tried. Various upward and downward door positions were used, to no avail. Cleats were then placed along the edges of the moveable doors (spawning the epithet "dragon's teeth") in order to disrupt the exhaust plume shear layer, with no consequence. A curtain of air was also blown between the doors and the front panel, but this too was ineffective.

The successful modifications were the addition of Helmholtz resonators to the sides of the EGMS conical diffuser and the introduction of water spray into the EGMS at the front panel opening. A Helmholtz resonator tuned to a chosen frequency acts as a band-pass filter for acoustic noise at that frequency. The 85 Hz tone was eliminated using this method. Water spray provides acoustic damping across the acoustic noise spectrum, and this method eliminated the 135 Hz tone.

Sverdrup has continued study of the nature and mechanism

of the EGMS discrete tone excitation. Contracts for experimental work have been given to the University of Tennessee Space Institute (UTSI) and the Georgia Tech Research Institute (GTRI), and theoretical analysis has been supported at Florida State University. These projects are in addition to analysis of the phenomenon by Sverdrup at AEDC. The effort described in this report falls into the latter category.

Method of Study

This study used a numerical methods approach to examine the acoustical properties of a vessel with characteristic dimensions comparable to the EGMS. The experimental studies at UTSI and GTRI were based upon reduced-scale models of an idealized EGMS. The idealized EGMS is shown in Fig. 3 with dimensions scaled to the nominal nozzle throat height. As a first approximation, the model vessel for this study was simplified to a right circular cylinder with length and

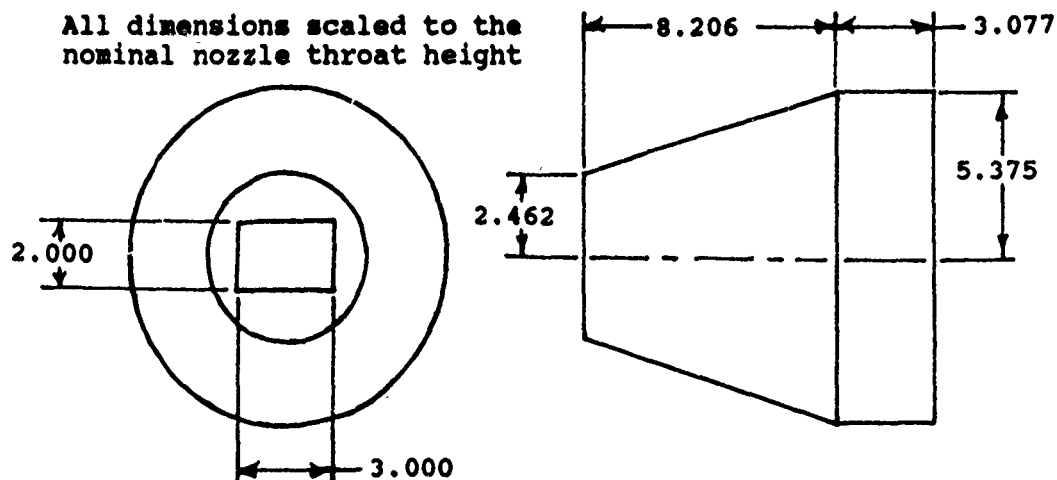


Fig. 3 - Idealized EGMS for use in acoustic analysis study

diameter equal to the idealized EGMS maximum diameter of 10.75 units.

The three main goals of this study were to demonstrate the acoustic modal response of the vessel, to demonstrate the use of Helmholtz resonators to reduce the acoustic response, and to seek physically more realistic models of the EGMS excitation mechanism.

The primary tool in pursuing these goals was the computer program BEMAP, developed at the University of Kentucky. This program was used on 80x86-family personal computers at Sverdrup. Both sound radiation from a vibrating surface and sound scattering due to incident sound waves can be calculated. The acoustic field inside or outside of an arbitrarily shaped body is determined by solving the acoustic wave equation using the boundary integral method.

The boundary integral method (also called the boundary element method) requires modeling only of the boundary enclosing (or enclosed by) the acoustic domain. Both geometry and boundary conditions must be specified. Unlike finite element or finite difference methods, the domain within the boundaries need not be modeled. Once the solution is found on the domain boundary, the solution at any point in the domain is easily obtained. This feature allows the program to run quickly enough for practical applications on a personal computer.

The surface geometry of a BEMAP boundary element model (BEM) is described by a set of Cartesian coordinates (node

points) grouped into either triangular or rectangular curvilinear units (elements). These triangular and rectangular elements are pieced together to form the surface of the BEM.

Elements may use either linear or quadratic shape functions, i.e. BEMAP will interpret the curvature of an element either linearly or quadratically. This shape function is also applied to the solution over the element. For linear triangular and rectangular elements, 3 and 4 node points are used, respectively. For quadratic triangular and rectangular elements, 6 and 8 node points are used, respectively. For comparably sized elements, quadratic shape functions offer greater accuracy than linear shape functions.

Preliminary Analysis

BEMAP requires specification of ambient and reference conditions. Values used for ambient density and speed of sound correspond to average values in the EGMS when the discrete tone problem arose. For the calculation of acoustic sound pressure levels, the standard reference acoustic pressure of 2×10^{-5} Pa was used.

For the analysis of problems with symmetric boundary conditions and geometry, BEMAP permits the use of half-models. This reduces modeling effort as well as CPU time. The geometry and modal responses of a circular cylinder and the idealized EGMS will have a plane of symmetry, so all BEMAP models used in this study are half-models.

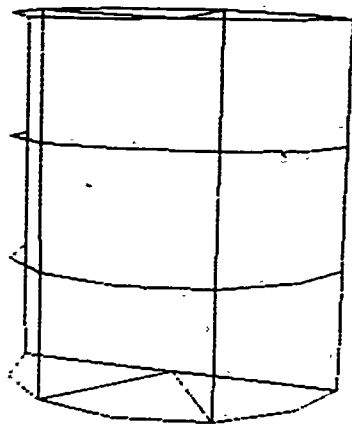
For acceptable solution accuracy, BEMAP requires approximately four linear elements per wavelength, or two quad-

ratic elements per wavelength. The frequency spectrum of interest ranges from 5 Hz to 250 Hz. Using quadratic elements would then require an element width of about 3.6 units. Since the model length and diameter are 10.75 units, this suggests using three elements along the length and diameter, and nine elements around the model circumference.

The version of BEMAP used at the beginning of this analysis limited the total number of available node points, which limited the number of elements. For this reason, coarse models (see Fig. 4) were created which used three elements along the model length, but only two elements across the model diameter, and six elements around the model circumference. Later, a more flexible BEMAP version was obtained, and fine models (see Fig. 5) were created for the cylinder and idealized EGMS with six elements along the model length, seven elements across the model diameter, and twelve elements around the model circumference.

Simple analytic solutions exist for the modal response of a right circular cylinder. For this case, the numerical solutions can be verified using the analytic solutions, thus lending credence to the BEMAP solutions for more complicated configurations. The analytic solutions are obtained by solving the acoustic wave equation with the separation of variables technique. The dependent variable is expressed as a function of distance along the length of the cylinder, radial distance from the centerline, angle around the circumference, and time. The standing wave solutions are then

Cylinder



EGMS

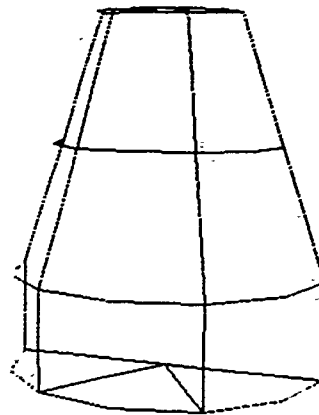
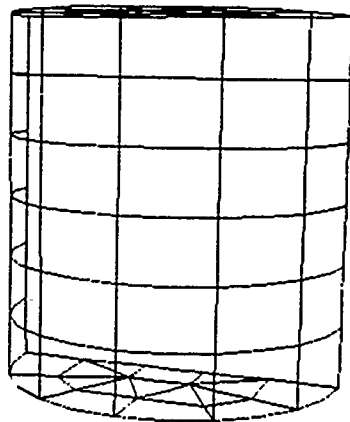


Fig. 4 - Coarse BEMAP models

Cylinder



EGMS

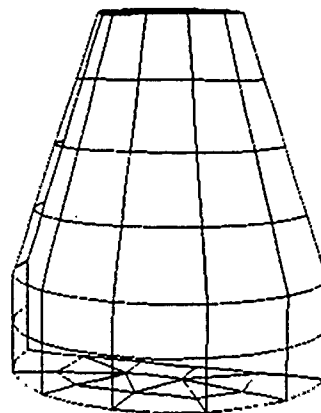


Fig. 5 - Fine BEMAP models

combinations of longitudinal waves, radial waves, and tangential (or circumferential) waves. A particular mode can be identified by the notation (m,n,k) , where the indices are the tangential, radial, and longitudinal mode numbers, respectively. A FORTRAN77 program called MODES was written to quickly calculate the normal mode frequencies for a right circular cylinder with specified dimensions and sound speed.

As part of this study, Helmholtz resonators were attached to the models. A Helmholtz resonator installed as a side-branch on an acoustic vessel is analogous to a mass-spring resonator attached to a simple beam. The mass-spring resonator is tuned to resonate at the same frequency as the beam, but with opposite phase, thereby absorbing energy from the oscillating beam. A Helmholtz resonator is a small acoustic volume chosen such that it resonates at the same frequency as a larger acoustic volume, but with opposite phase. When the smaller volume is attached to the larger volume, acoustic energy tends to be cancelled at the tuned frequency. A FORTRAN77 program called HRDESIGN was written to quickly calculate the dimensions of a Helmholtz resonator for a specified frequency and sound speed.

Acoustic Modal Responses for Cylinder

The acoustic response inside a right circular cylinder with closed ends was examined in the range 5 Hz to 250 Hz. The cylinder had a length and diameter of 10.75 units. An acoustic point source was placed inside the cylinder at a position 1.0 unit from one end-wall and 1.0 unit radially

from the cylinder centerline. The unit sound power point source was positioned on the plane of symmetry of the cylinder half-model.

The longitudinal modes for this case were examined with the coarse model. The analytic solution predicts the $(0,0,1)$ and $(0,0,2)$ modes will occur at 66.8 Hz and 133.7 Hz, respectively. Over the range 5 Hz to 250 Hz with a 5 Hz frequency increment, the numerical solution showed that average sound pressure inside the cylinder had spikes at 65 Hz and 135 Hz. The range 50 Hz to 150 Hz with a 1 Hz increment was then examined. These results displayed average cylinder pressure spikes at 67 Hz and 134 Hz. The pressure distribution on the surface of the cylinder for 67 Hz and 134 Hz (see Fig. 6) confirms that these are the first two longitudinal modes. For color figures in this report, red and indigo correspond to high and low values, respectively.

The fine model showed similar results for the longitudinal modes. Average pressure inside the cylinder showed spikes at 67 Hz and 134 Hz. Surface pressure for the fine cylinder model (see Fig. 7) compares well with results from the coarse model for the $(0,0,1)$ mode. For the $(0,0,2)$ mode, the fine model more robustly displays the solution. It can be concluded that the coarse model is adequate for analysis of the $(0,0,1)$ mode, but the fine model may be more appropriate for analysis of the $(0,0,2)$ mode.

The tangential modes were then investigated for the coarse model. The analytic solution predicts the $(1,0,0)$

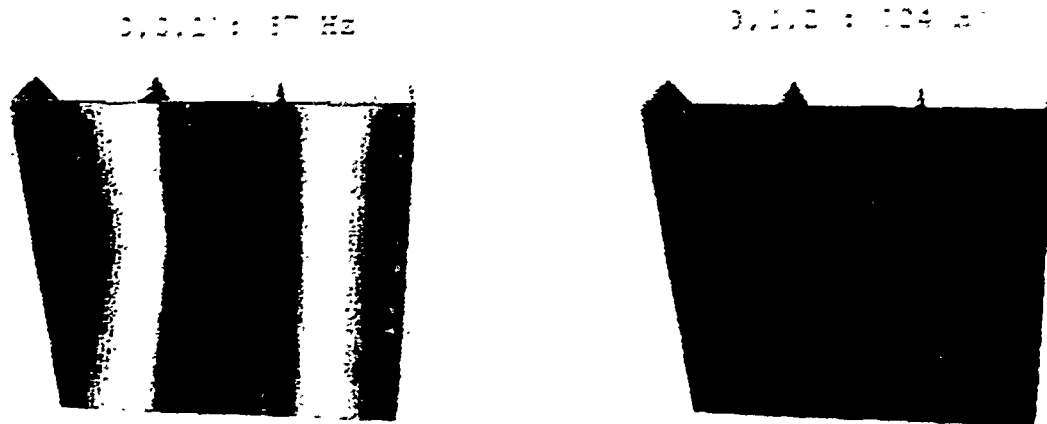


Fig. 6 - Pressure amplitude for longitudinal modes of the coarse model cylinder

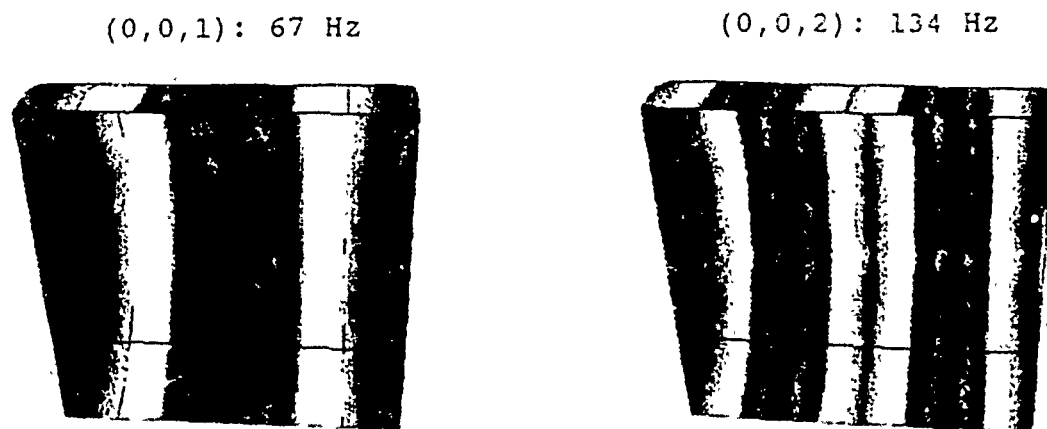


Fig. 7 - Pressure amplitude for longitudinal modes of the fine model cylinder

and $(2,0,0)$ modes will occur at 78.3 Hz and 130.0 Hz, respectively. Over the range 5 Hz to 250 Hz with a 5 Hz frequency increment, the numerical solution showed a spike in average sound pressure at 80 Hz, but none at 130 Hz. The range 50 Hz to 150 Hz with a 1 Hz increment was then examined, and displayed an average pressure spike at 79 Hz, but still none at 130 Hz. The surface pressure distribution for 79 Hz (see Fig. 8) confirms that this is the first tangential mode.

The fine model displayed average pressure spikes at 78 Hz and 130 Hz. Surface pressure distributions for 78 Hz and 130 Hz (see Fig. 9) confirm that they are the $(1,0,0)$ and $(2,0,0)$ modes, respectively. It can be concluded that the coarse model may be adequate for analysis of the $(1,0,0)$ mode, but the fine model must be used for higher tangential modes.

The radial modes were finally examined for the coarse model. The analytic solution predicts the $(0,1,0)$ and $(0,2,0)$ modes will occur at 163.1 Hz and 298.5 Hz, respectively. The frequency for the $(0,2,0)$ mode falls outside the range of interest for this study, so only the first radial mode was sought. Over the range 5 Hz to 250 Hz with a 5 Hz frequency increment, the numerical solution showed a spike in average sound pressure at 165 Hz. The range 163 Hz to 166 Hz with a 1 Hz increment was then examined, and displayed an average pressure spike at 164 Hz. In Fig. 10, the surface pressure distribution for 164 Hz confirms that this

$(1,0,0)$: 79 Hz

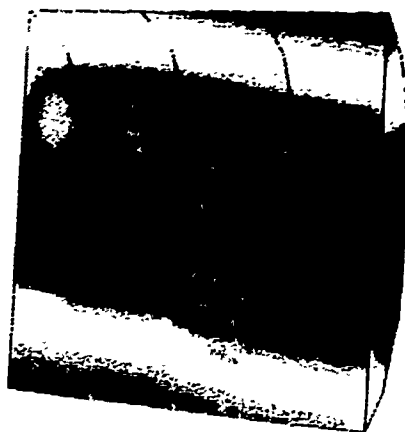
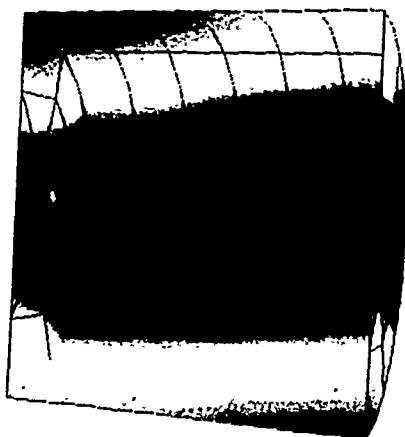


Fig. 8 - Pressure amplitude for tangential mode of the coarse model cylinder

$(1,0,0)$: 78 Hz



$(2,0,0)$: 130 Hz

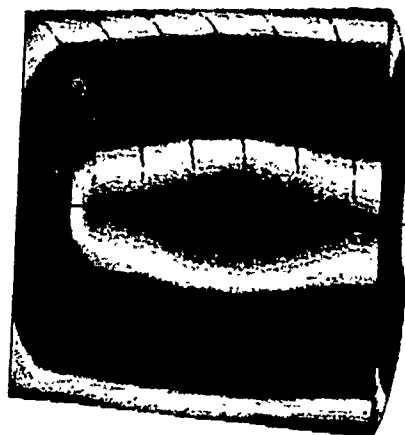


Fig. 9 - Pressure amplitude for tangential modes of the fine model cylinder

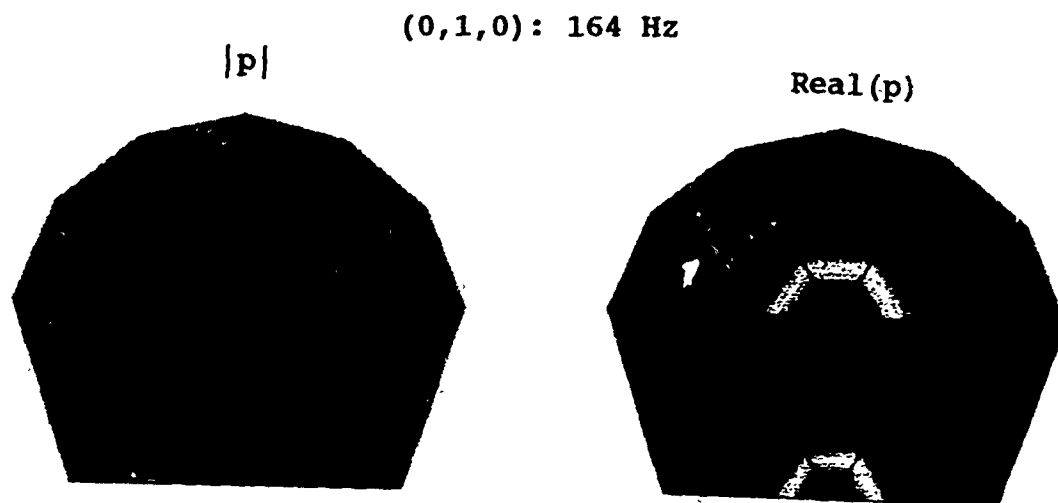


Fig. 10 - Pressure values for radial $(0,1,0)$ mode of the coarse model cylinder

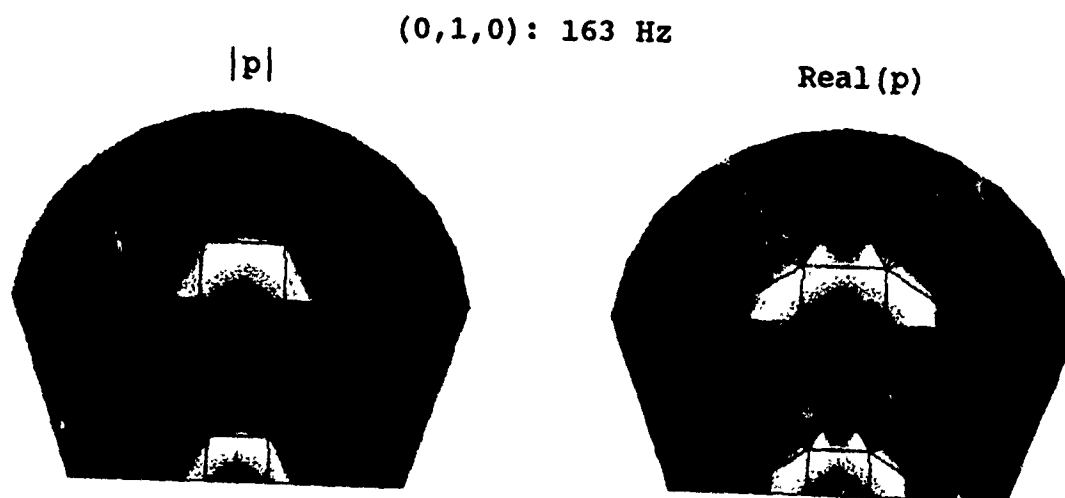


Fig. 11 - Pressure values for radial $(0,1,0)$ mode of the fine model cylinder

is the first radial mode. To increase the clarity of the radial mode display, both pressure amplitude and the real part of pressure are shown.

The fine model showed a similar result for the (0,1,0) radial mode. Average pressure inside the cylinder showed a spike at 163 Hz. Surface pressure for the fine cylinder model (see Fig. 11) shows that the solution is more robust than for the coarse model. It can be concluded that the fine model is probably more desirable for analysis of the radial modes.

The results obtained for the circular cylinder model lead to the not surprising conclusion that the fine model gives better solutions than the coarse model. Nevertheless, the coarse model performs fairly well for the fundamental modes. The coarse model can, at the least, serve to qualitatively predict trends at these lower frequencies. The fine model should be used for more quantitative work. For the fine cylinder model the average sound pressure level spectrum, from 5 Hz to 250 Hz with a 1 Hz increment, is displayed in Fig. 12. The frequency values shown for each spike agree with the analytic solutions to the number of significant figures, except for the 187 Hz value which is 1 Hz too high. For comparison, Fig. 13 shows the average sound pressure level spectrum for the coarse model.

Conclusion

This study has established that BEMAP correctly determines the normal mode acoustic responses in a circular cyl-

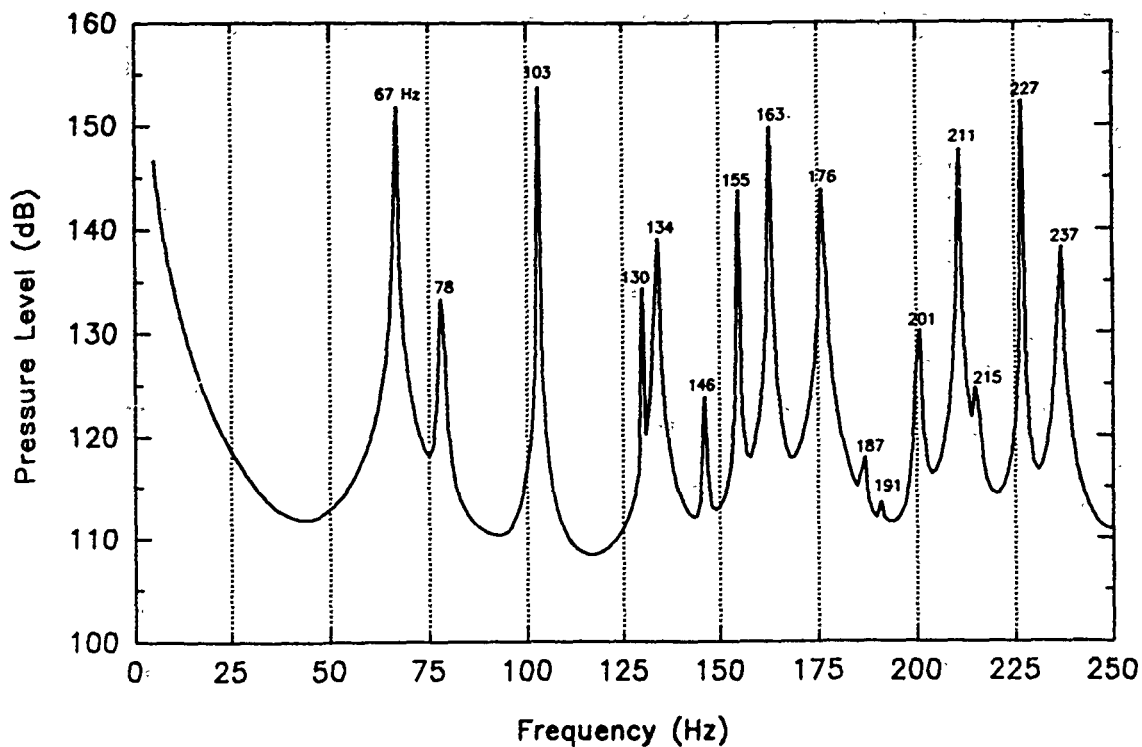


Fig. 12 - Average acoustic sound pressure level inside the fine model cylinder

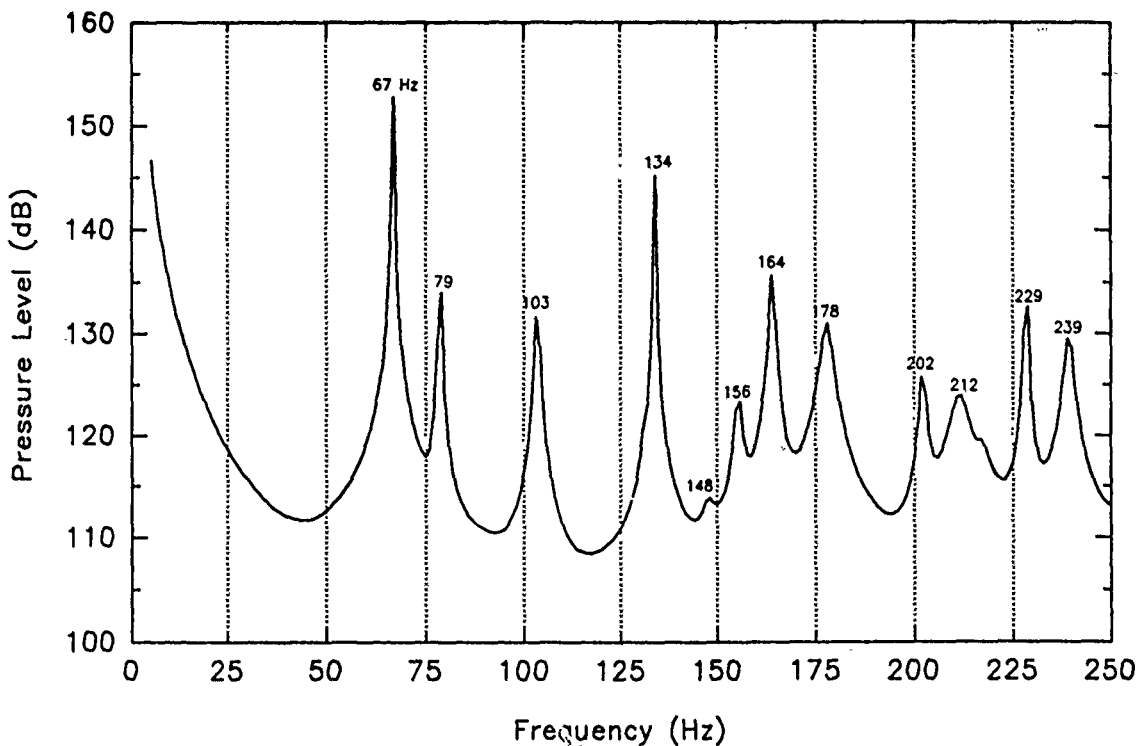


Fig. 13 - Average acoustic sound pressure level inside the coarse model cylinder

inder. The effects of boundary element model resolution were examined for the longitudinal, tangential, and radial modes of a circular cylinder approximating an EGMS. It was concluded that the coarse resolution model is appropriate for qualitative work, but the fine resolution model should be used for quantitative analysis.

Space does not permit adequate description of other results obtained in the course of this ten week project. Other topics covered were examination of the use of Helmholtz resonators, acoustic response of the idealized EGMS, investigation of physically more realistic model geometries for the EGMS, and their acoustic responses.

This work represents a beginning for the use of the boundary element method in the analysis of the Exhaust Gas Management System. The foundations laid here will be built upon through continuation work conducted by Dr. Jones at Sverdrup.

References

Beranek, L. L., ed., Noise and Vibration Control, McGraw-Hill, 1971.

Kinsler, L. E., Frey, A. R., et al., Fundamentals of Acoustics, Wiley, 1982.

Morfey, C. L., "Rotating Pressure Patterns in Ducts: Their Generation and Transmission", J. Sound Vib. Vol. 1, pp. 60-87, 1964.

Seybert, A. F., Wu, T. W., and Wan, G. C., User's Manual Computer Program BEMAP Version 2.43, Spectronics Inc., September, 1990.

Tyler, J. M., and Sofrin, T. G., "Axial Flow Compressor Noise Studies", SAE Transactions Vol. 70, 1962.

1990 USAF-RFL SUMMER FACULTY RESEARCH PROGRAM/

GRADUATE STUDENT RESEARCH PROGRAM

**Conducted by the
Research and Development Laboratories**

FINAL REPORT

**THE INFLUENCE OF MATERIAL PROPERTIES AND SHAPE ON THE
MEASURED HEAT-FLUX DISTRIBUTION USING A HEMISPHERICAL
COAXIAL THERMAL COUPLE TEMPERATURE PROBE**

| | |
|---------------------------|---|
| Prepared by: | William Glenn Wilk |
| Academic Rank: | Graduate Student |
| Department and | Aerospace Engineering |
| University | The University of Alabama |
| Research Location: | Arnold Engineering Development Center Arnold AFB, TN 37389 |
| USAF Researcher | Dwayne Carver, Calspan Corporation AEDC Division |
| Date: | August 23, 1991 |

**THE INFLUENCE OF MATERIAL PROPERTIES AND SHAPE ON THE
MEASURED HEAT-FLUX DISTRIBUTION USING A HEMISPHERICAL
COAXIAL THERMAL COUPLE TEMPERATURE PROBE**

William Glenn Wilk
Graduate Student, The University of Alabama

ABSTRACT

The primary objective of the project was to evaluate the effect of varying materials, their properties, and the hemispherical shape on a coaxial surface thermocouple heat flux probe. The results of the evaluation were used to develop a more precise heat flux reduction scheme. The present data reduction program to calculate the heat flux of a flow field assumes the probe behaves as a semi-infinite body of a single material at standard temperatures. Using the temperature time histories of a probe passing through the field under these assumptions results in an unsymmetrical heat-flux distribution whose magnitude is too low. A finite-element model of the probe was used with the computer code, TRAX, to perform this evaluation. After thorough investigation, the reduction procedure, COAXVP, was modified to include material property effects for the complete temperature range. Further improvements could be achieved by modifying future probe's shape (increasing nose radii) to closely match the semi-infinite body.

Acknowledgments

I wish to thank Mr. Dwayne Carver from Arvin Calspan Corporation for giving me guidance and support in achieving the project objectives. Finally, acknowledgment must be given to the Air Force System Command and the Air Force Office of Scientific Research for providing the research opportunity through the administrative support of Research and Development Laboratories.

INTRODUCTION

The extremely high temperature of the flow in a high-enthalpy arc-heated tunnel prevents the use of available heat-flux transducers directly. Instead, the heat flux is found indirectly using stagnation temperature probes that are quickly sweep through the flow field(Fig. 1). The temperature time history is recorded and then numerically integrated to determine the heat flux. The measuring probes used are the null-point calorimeter and the Chromel-constantan coaxial surface thermocouple. (Fig. 2)(Ref. 1)

Current test results show that the coaxial surface thermocouple (supported in a stainless steel hemispherical body) measures lower temperatures after initial injection and therefore, lower heat-flux values than expected and seen using the copper bodied null-point calorimeter. Furthermore, the heat-flux profile is unsymmetrical. The calculated heat-flux value is consistently higher on the probe entrance side and drops steadily until the exit side (Graph 1).

The influence of the hemispherical body shape, the difference in the probe materials, and the variation of the material properties with temperature are not factored into the current integration reduction program of the coaxial thermocouple. By accounting for these influences in a modified reduction program, a more accurate heat-flux distribution can be found.

APPROACH

To verify that the shape and the material properties are an influence and to, therefore, produce a proper heat-flux reduction program, a finite-element model was created of the stagnation or tip portion of the coaxial surface thermocouple(Fig 3). The model is used with the heat conduction finite-element code, TRAX(Ref. 2). Boundary conditions of heat flux and heat transfer coefficient can be specified between any two element nodes on the model. Multiple time intervals with different boundary conditions and updated material thermal properties may be specified. TRAX outputs the temperature time history for each nodal point in the model. The temperature time history of the node corresponding to the coaxial thermocouple junction was used as the input to the integration-reduction code, COAXVP. The reduction program assumes the inputted temperature time history, $T(t)$, of the probe as that of a homogeneous semi-infinite body. The numerical integration equation for heat flux is:

$$\dot{q}(t) = \frac{2(\rho C_p K)^{1/2}}{\pi^{1/2}} \sum_{i=1}^n \frac{T_i - T_{i-1}}{(t_n - t_i)^{1/2} + (t_n - t_{i-1})^{1/2}} \quad (1)$$

where ρ , C_p , and K are the density, specific heat, and thermal conductivity(Ref. 1).

RESULTS

Various cases were performed to verify the TRAX model and to establish an understanding of the hemispherical probe behavior. The initial case, restricted to the single material - constantan, was subjected to a steady, constant heat flux. The material properties were limited to room temperature. The model shape and temperature time history were confirmed with an analytical solution of a constant heat flux on a homogeneous sphere(Ref. 3) Both cases indicated that the heat-flux value should rise across the flow field due to the small spherical radius of only 0.25 inches(Graph 2). This suggested that the radius was too small to continue assuming a semi-infinite body. However in later cases, a more accurate representation of the heat flux distribution on the hemispherical probe was used. The heat flux at the stagnation point was equal to that of the flow field. The remaining heat-flux boundary conditions decreased according to Lee's Distribution of a sphere in a flow of constant heat flux. As seen in Graph 3, this distribution along the boundary negated the nonsemi-infinite, small radius effects. The heat flux rises only slightly. Accounting for variable properties of the different materials of the sensor, reduced the heat-flux distribution across the flow field as seen with the actual coaxial surface temperature thermocouple data(Graph 4).

As a result, the integration, reduction program, COAXVP, was modified to include property effects by allowing Beta, $\sqrt{\rho CK}$, to be a function of temperature. The calculated heat-flux magnitude and distribution using experimental data was comparable to that of the null-point calorimeter (Graph 5)

CONCLUSION

By accounting for variable properties, the heat-flux values and distributions should be a more accurate representation of the flow field especially at the higher temperatures associated with a high-enthalpy arc-heated wind tunnel. Although the overall effect of the shape of the probe can be neglected, future probes with a larger radius of approximately one inch would behave more as a semi-infinite solid. This could improve heat-flux calculations at lower temperatures.

REFERENCES

1. Carver, D.B. and Kidd, C.T., "Heat-Transfer Measurement Uncertainty In Arc-Heated Flows," Proceedings of the 37th International Instrumentation Symposium, ISA Paper #91-115, 1991, pp.951-964.
2. Kidd, C.T., "Lateral Heat Conduction Effects on Heat-Transfer Measurements with the Thin-Skin Technique," ISA Transactions, Vol 26, No.3.
3. Carslaw, H.S. and Jaeger, J.C., Conduction of Heat in Solids, Clarendon Press, Oxford, 1959.

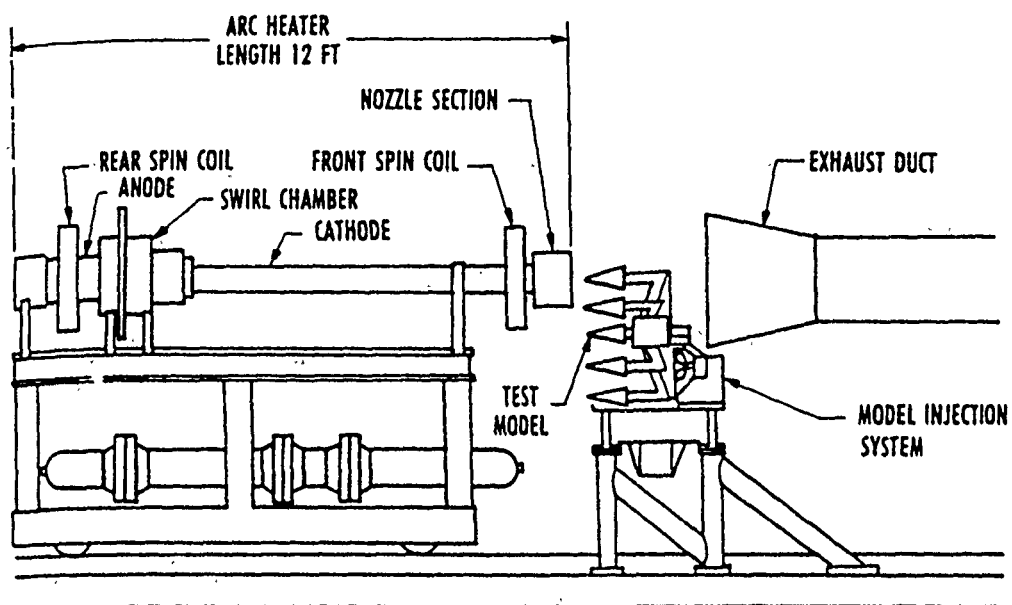
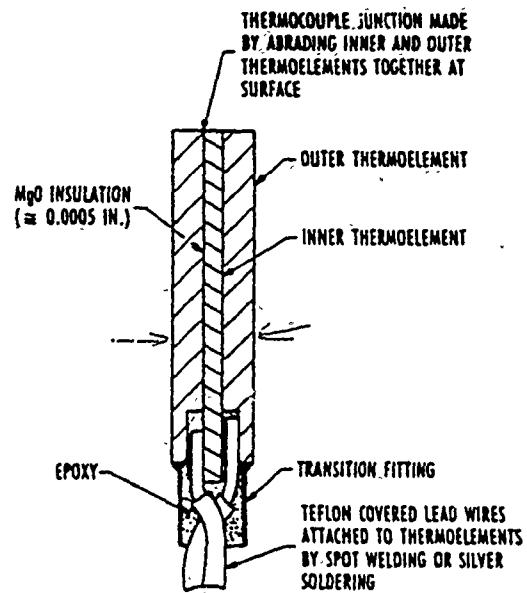
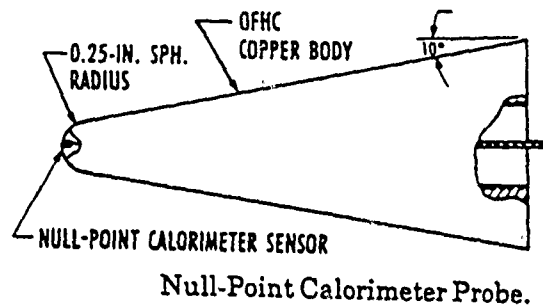


Figure 1. Heat-HR Test Unit Components. (Ref 1)



Sketch of Coaxial Surface Thermocouple, No Scale.



Null-Point Calorimeter Probe.

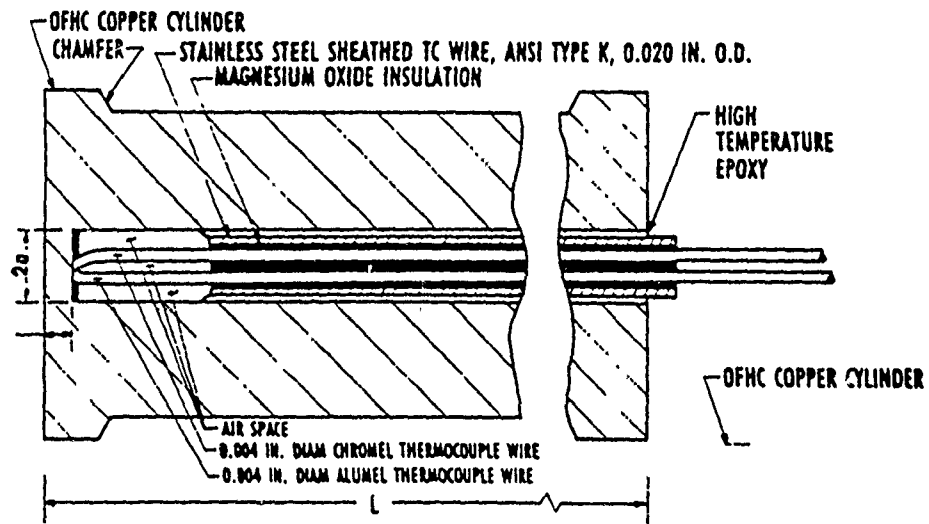


Figure 2. Section View of Null-Point Calorimeter Assembly, No Scale. (Ref 1)

TRAX MODEL GEOMETRY

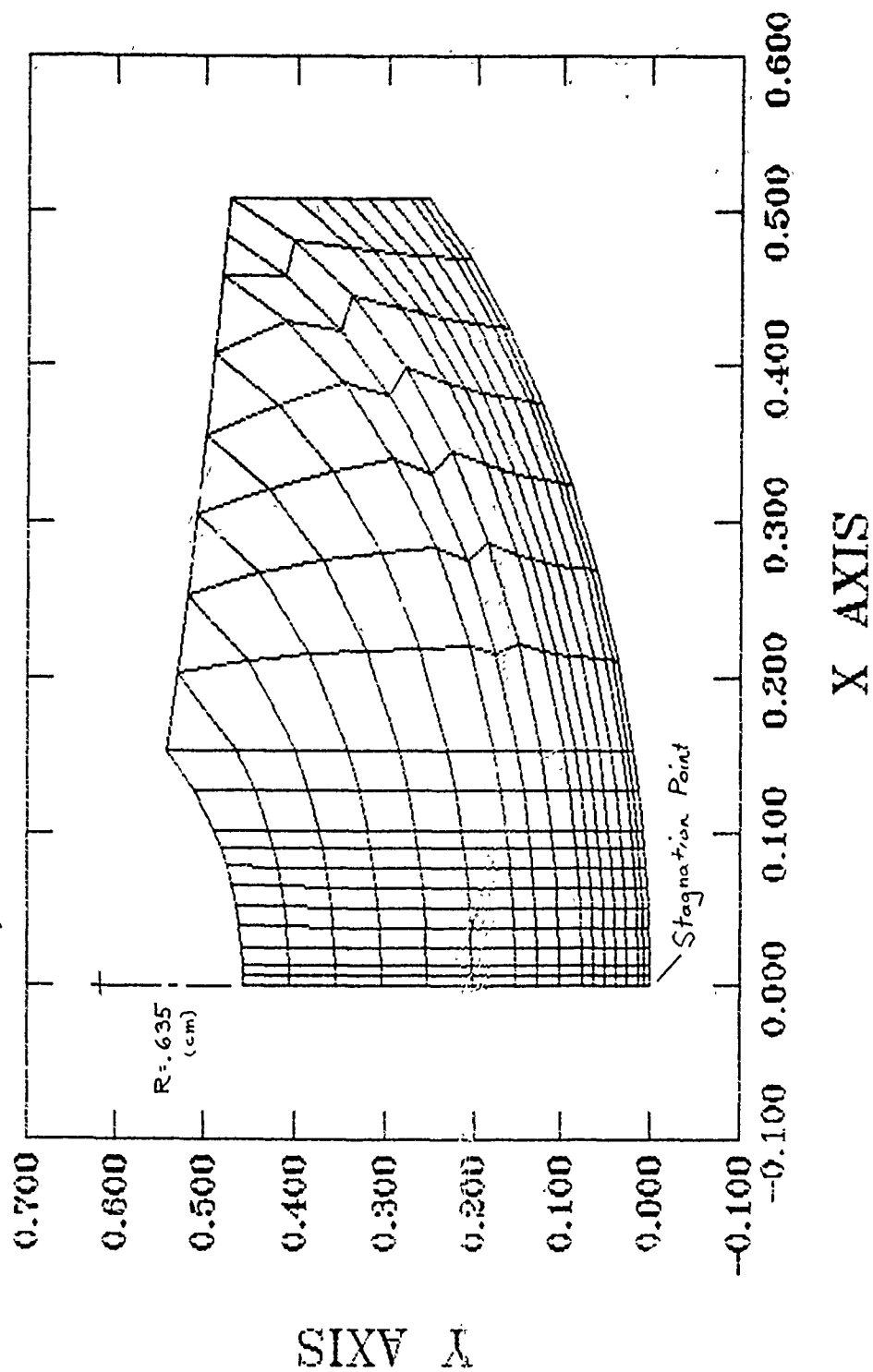


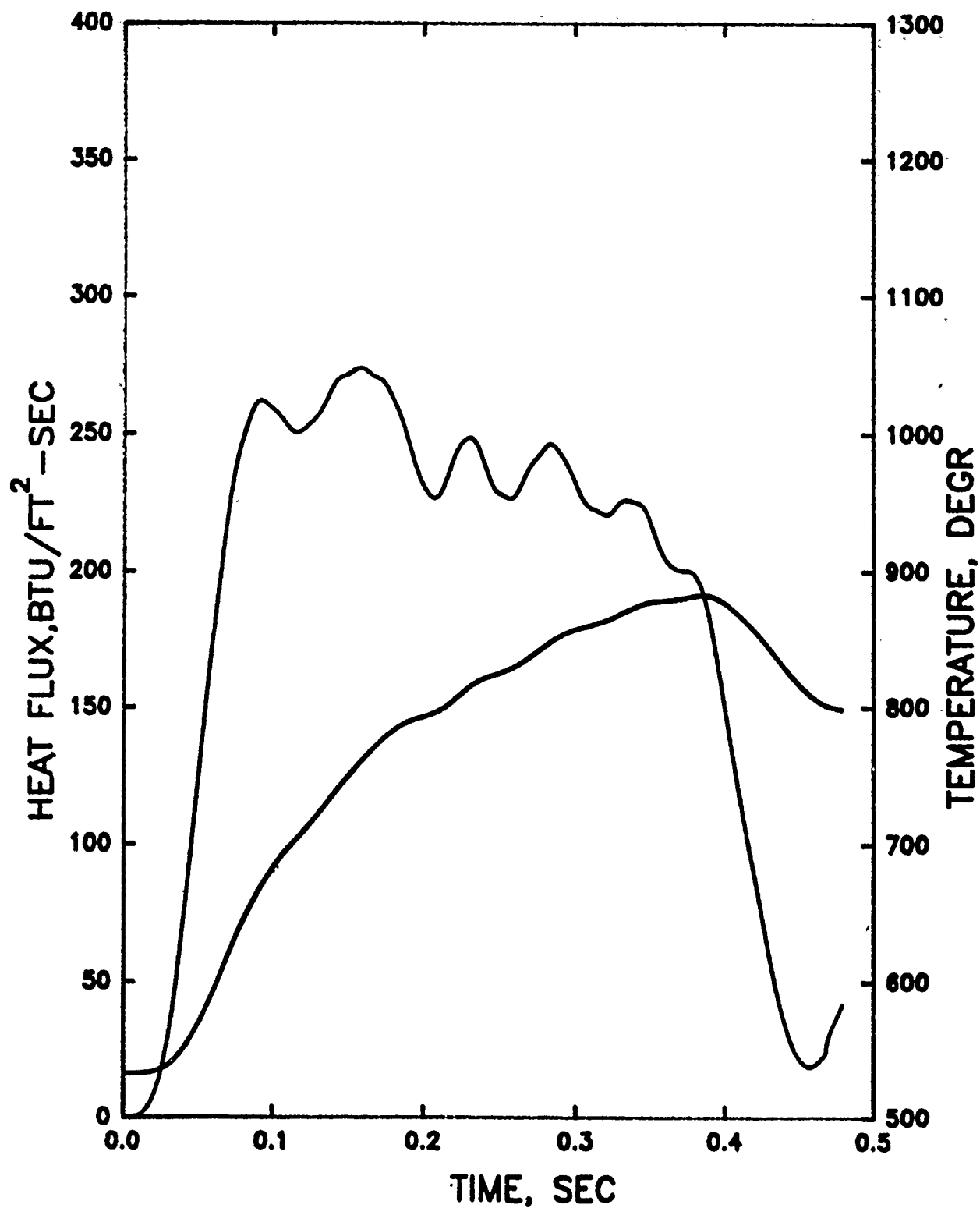
Fig 3

H2 CALIBRATION DATA CAL 1S RUN 6 SWP1

BETA0 = 0.425

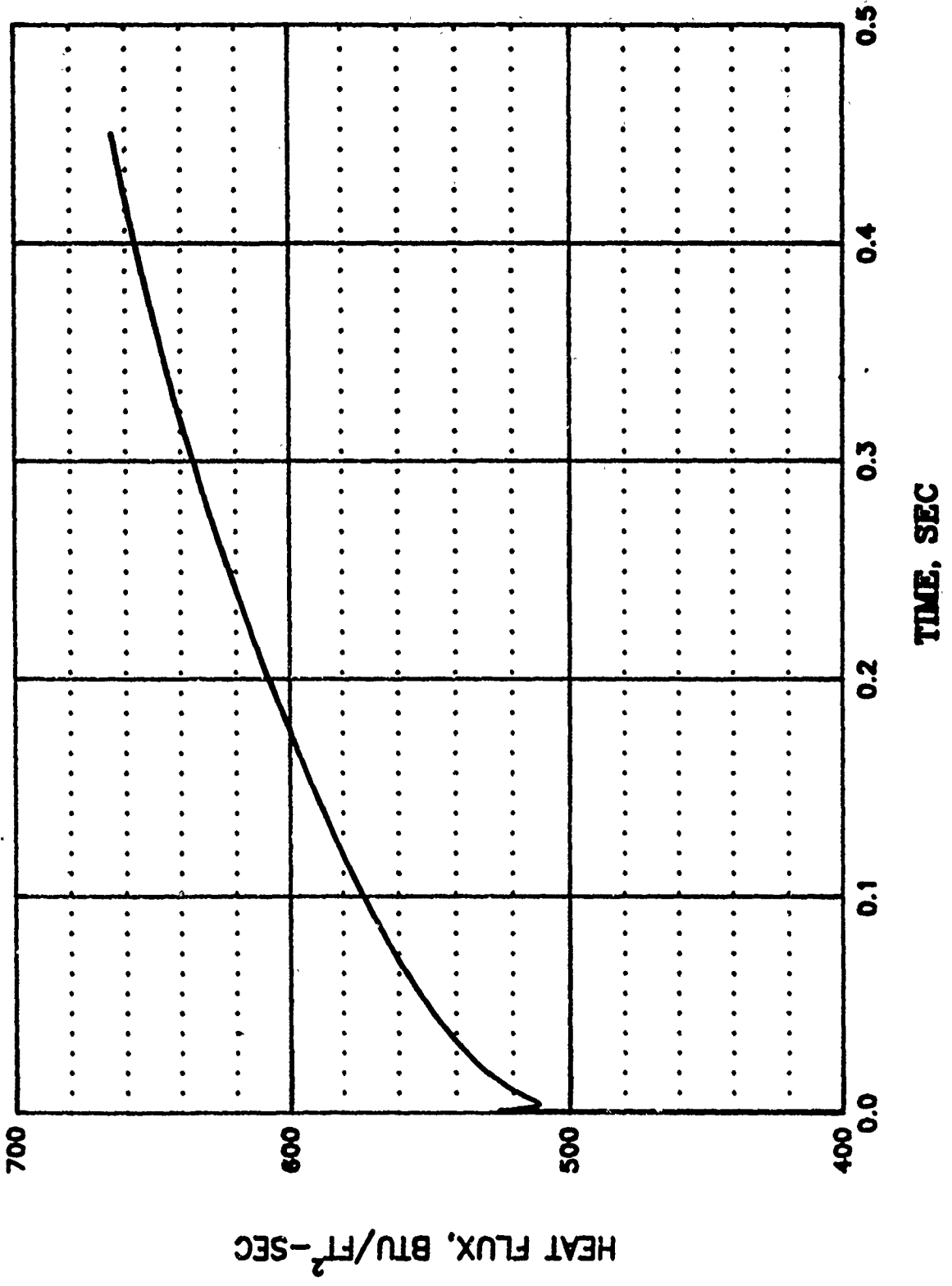
BETA1 = 0

BETA2 = 0



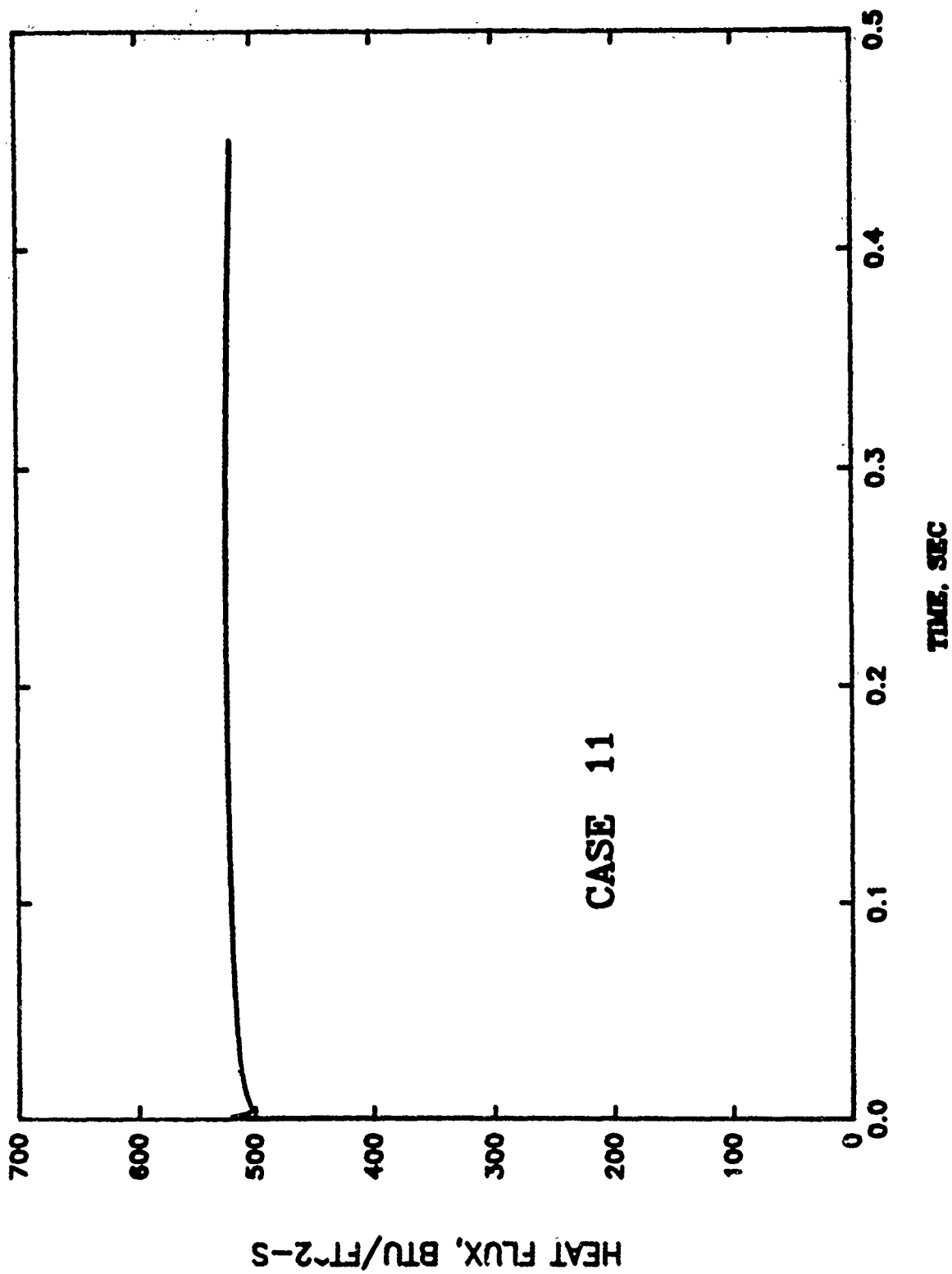
Graph 1

TRAX SOLUTIONS INPUT TO COAXVP CASE 1

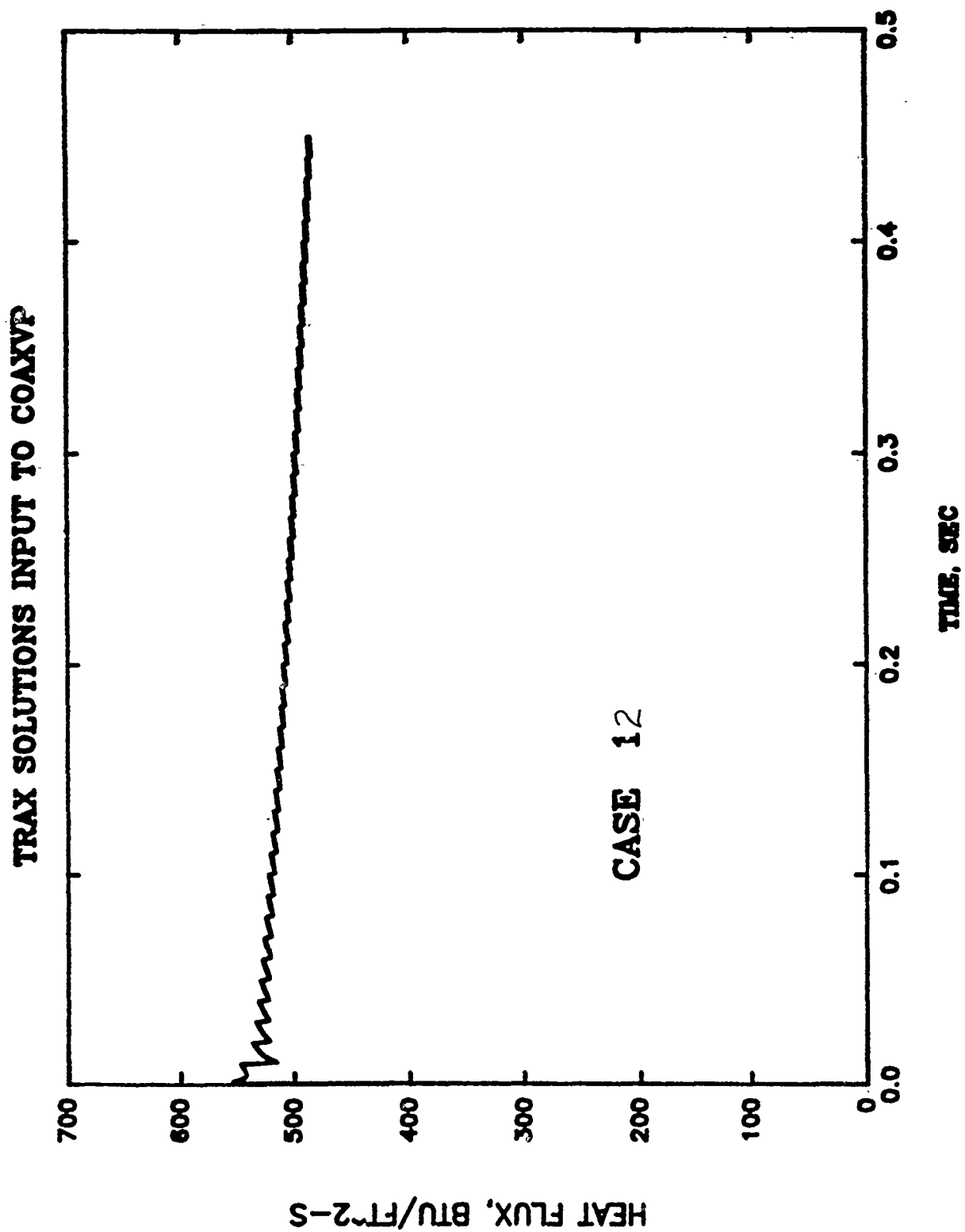


Graph 2

TRAX SOLUTIONS INPUT TO COAXVP



Graph 3



Graph 4

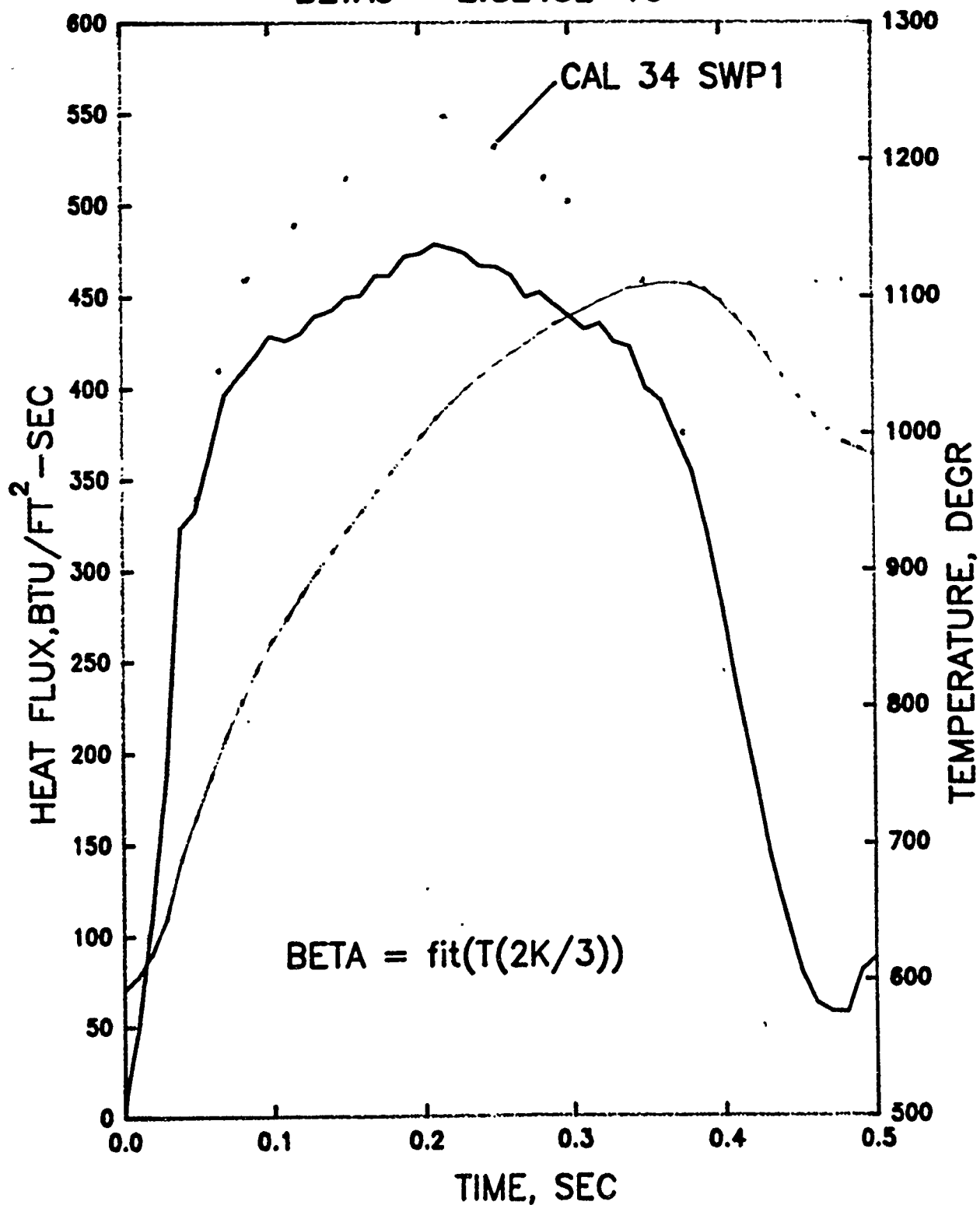
H2 CALIBRATION DATA CAL 1S RUN 5 SWP8

BETA0 = 0.18729

BETA1 = 7.2426E-4

BETA2 = -5.747E-7

BETA3 = 2.3245E-10



Graph 5

32-17

The Effects of a Passive Porous Surface on Flow Field Development Under both Static and Dynamic Conditions

Gregory A. Addington
Master Degree Candidate
University of Colorado at Boulder

Abstract

A passive porous suction surface was added to a NACA 0015 wing of otherwise conventional construction. Static surface pressure measurements indicate that the adverse pressure gradient usually present over the aft portions of such a wing were attenuated by the addition of this porous surface. Pressure gradients near the leading edge, however, are maintained in magnitude although local C_p magnitudes are generally less than those collected for non-porous wing configurations. When this pressure data is combined with flow visualization, a separated region is evident over the wing's aft portions even low angles of attack. Flow visualization taken during constant rate ramp motions in pitch indicate that no clearly evident and repeatable dynamic stall vortex is produced over the porous surface at low pitch rates, whereas one is evident for the non-porous and solid surface wings. As non-dimensional pitch rates are increased, a dynamic stall vortex becomes evident, although initiating earlier in the pitch motion than for non-porous or solid wings.

Introduction

Porous or semi-porous surface wings and airfoil and wing surfaces have long been used to actively or passively affect changes in the pressure and flow fields elicited by the wing's movement through air. Some of these past studies have used active suction or blowing to affect flow field or boundary layer control^{1,2}, including a number of flight test programs^{3,4}. Others have used the pressure differences that exist across cavities or surfaces to passively alter flow characteristics. Such studies include Chen and colleagues⁵, and other's^{6,7} use of porous surfaces to reduce wave drag; Soderman's use of perforated surfaces to control wind tunnel resonances⁸; and Abbott's use of multi-slotted wings for laminar flow control⁹.

Because the process of dynamic stall is characterized by a large variation of pressures across the suction surface¹⁰⁻¹², dynamic stall pressure fields would be dramatically altered by the addition of a passively forced

porous suction surface. This pressure field attenuation would be accomplished in much the same manner as was demonstrated by Abbott²: lower pressure at the leading edge would create an "internal circulation" within the underlying cavity and the suction surface, thus affecting flow reversal within the boundary layer and, therefore, the vorticity distribution over the suction surface. It was intended that by radically altering the surface pressure experienced by the dynamically pitching wing, the relative importance of boundary layer development and overriding flows could be determined.

To provide a baseline from which to start this study, static angle of attack surface pressure and flow visualization data collection is a part of the current study. To date, only flow visualizations of flow fields elicited by dynamic wing motions have been accomplished.

Nomenclature

| | |
|--------------|--|
| DSV | Dynamic Stall Vortex |
| α | angle of attack (degrees) |
| α^+ | non-dimensional pitch rate ($d\alpha/dt * c / U_{inf}$) |
| $d\alpha/dt$ | pitch rate (rad/sec) |
| c | chord length (ft) |
| U_{inf} | freestream velocity (ft/sec) |
| t^* | non-dimensional time ($t * U_{inf} / c$) |
| C_p | pressure coefficient ($(p_{local} - p_{inf}) / q_{inf}$) |

Methods

Experiments were conducted in the Frank J. Seiler Research Laboratory (FJSRL) and Department of Aeronautics (DFAN) 3 ft x 3 ft low speed wind tunnels (LSWT) located at the United States Air Force Academy. These wind tunnels are of a nearly identical design, allowing for trouble-free transition and highly comparable results from one wind tunnel to the other. Static pressure measurements were made in the DFAN LSWT while all flow visualizations were performed in the FJSRL LSWT.

For this experiment, a porous suction surface wing was constructed from hollow extruded aluminum airfoil stock of 0.5 ft chord length and 2 ft in span. One wingtip was fitted with a solid end cap cut to the same airfoil cross section while the other was fit with a 1 ft diameter circular splitter plate and mountings for wind tunnel mechanisms. The splitter plate was constructed of 0.25 in thick aluminum plate machined with a sharp perimeter, and was centered on the wing's midchord. 1.5 ft of one surface

was removed and replaced by one of two 26 gage perforated brass sheets (see Table 1 for brass sheet specifications). As can be seen in Fig. 1, this construction created a hollow cavity over which laid the brass sheet. The two outlying ribs within this cavity were added to give structural rigidity required for dynamic motions and to give the brass sheet form. A center rib containing 13 pressure ports was constructed so that their inlets would be flush with the outer surface of the brass sheet. These pressure ports were placed as shown in Fig. 2.

During all aspects of the tests, the wing-splitter plate combination was mounted to a 1.125 in diameter steel shaft. This shaft was connected to a Superior Electric NH172 synchronous stepper motor via a gear linkage with a 4:1 reduction ratio. The use of this mechanism allowed for the attainment of highly repeatable and accurate constant rate pitch motions about the quarter chord as well as accurately measured static angles of attack. The wing's motions were controlled by embedded software on a Zenith 248 microcomputer during the constant rate motions, and a MASSCOMP 5500 for the static tests.

Pressure data was taken using an Endevco 8507C-2 miniature pressure transducer. For static α surface pressure measurements, the pressure ports were plumbed to a Scanivalve model J stepper motor-switched manifold system, with the pressure transducer at the outlet of the manifold. The Scanivalve switching and data acquisition were both controlled by a Keithley model 500 data acquisition system, which was in turn controlled by BASIC software embedded on a Zenith 248 PC. Pressure transducer signals for this phase of the study were passed through a low pass filter (bandwidth of 1 Hz) and then amplified by a gain of 1000 before digitization by the Keithley and collection by the Zenith. Ensemble averaging of 1000 data points taken at 1 kHz and simple computations gave local C_p values.

Flow visualizations were collected using the same wing-splitter plate combination. A tungsten wire 0.005 in. in diameter was placed approximately two chordlengths ahead of the wing orthogonal to the midspan location. The smokewire was lightly coated with SAE 30 weight motor oil, which was allowed to form discrete beads. The smokewire was then ohmically heated, vaporizing the oil, thus producing a dense, homogeneous sheet of streaklines.

The visualized flow field was illuminated by two Berkley Photo Colortran tungsten-halogen lamps. Illumination intensity was controlled by

a 3.1 kVA variac. Flow visualization images were collected at 250 frames per second by a Kodak Ektapro 1000 high speed video imaging system equipped with a Kodak image intensifying camera system. The camera was fitted with a Fujinon telephoto zoom lens set to an f-stop of 2.5 and focal length of approximately 50 millimeters. Imager gate time was 500 microseconds while gain was varied as circumstances warranted.

Results

A comprehensive static angle of attack study was undertaken as a part of this experiment to first gain insight into the characteristics of the wing-perforated screen mating. This study was conducted in two parts: first, the larger open area sheet was backed with a thick paper matting, thus giving an approximately solid surface (termed the non-porous surface) wing with the surface roughness of the sheet attached. The paper backing was then removed and the full porosity was used. Finally, the large open area sheet was replaced with the lower open area sheet and the procedures were again repeated. The small open area sheet was not used in the non-porous configuration during the surface pressure measurements.

Static α explored began at 0° and increased by 1° increments to 20° , followed by 25° , 30° , 40° , 50° and 60° . The freestream velocity used during pressure measurements was 30 ft/sec, yielding a Reynolds number of 72000. For flow visualizations, U_{inf} was reduced to 20 ft/sec (Reynolds number = 48000) in order to preserve smoke particle density. For more detailed analysis, flow visualization frames taken at 0° through 30° in increments of 5° , then 40° , 50° and 60° were selected, portions of which are included in this report.

C_p data for the non-porous condition with the large hole open area sheet attached (Fig. 3) shows that a prominent suction peak was formed in the vicinity of the LE, followed by a slow tapering pressure recovery area aft of 10% c at an α of 0° . Theoretical data¹³ for a solid surfaced airfoil at $0^\circ \alpha$ is included in Fig. 3. Flow visualizations at this α (Fig. 4) indicate that the flow remained attached over the entire chord length. As α was increased, the magnitude of the suction peak increased, culminating at a value of -3.1 at approximately $14^\circ \alpha$. Beyond an α of 15° , the leading edge C_p values lessened in magnitude, and became nearly homogeneous with the remainder of the suction surface by 30° . Beyond $30^\circ \alpha$, C_p values simultaneously decreased from approximately -0.7 to approximately -0.95 at

$60^\circ \alpha$. Corresponding flow visualizations (Fig. 4) indicate that trailing edge separation began between 5° and $10^\circ \alpha$, that a highly turbulent region resided over the suction surface at 15° , and that flow has completely separated by 20° . Jacobs and Abbott¹⁴ indicate that a NACA 0015 with solid, polished surfaces experiences lift stall at an α of approximately 11° at a Reynolds number of 84000. Flow visualizations for the small open area sheet with backing at static angles of attack (not included) show only minimal differences with the large open area sheet.

The graphs of C_p vs c in Fig. 3 also show an apparent shift in stagnation point away from the leading edge through the increasingly negative values of C_p at the leading edge at values of α below 15° . This stagnation point shift away from the leading edge along the pressure surface as α increases is confirmed by flow visualizations and anticipated by theory^{15,16}.

Removal of the paper backing caused a decrease in the magnitude of C_p ahead of $c/4$ (Fig. 3) compared to the non-porous condition at all α less than 20° , although the pressure gradients in this region did remain of the same magnitude (see Fig. 5). The minimum value of C_p encountered for this configuration was -0.87 at an α of 18° . Furthermore, aft of $c/4$ and at these same α 's, no appreciable pressure recovery is seen in the C_p data; instead, a nearly isobaric region is observed. Flow visualizations for $0^\circ \alpha$ (Fig. 6) show a separated region in which an energetic shear layer existed aft of approximately $c/4$. As α was increased, this shear layer increased in width (see Fig. 8). C_p values for the non-porous and this porous condition began to converge past an α of 15° , and finally converged at $25^\circ \alpha$, after which there are no appreciable differences in pressure profiles between the two configurations. Like for the non-porous condition, an apparent stagnation point shift did occur, but leading edge C_p values did not reach as great a magnitude as for the non-porous condition.

Changing to the lower open area sheet, the most notable difference between the two porosities at $0^\circ \alpha$ was a slight suction spike (Fig. 3) ahead of $5\% c$, after which C_p values level off to near those of the larger porosity wing configuration. The suction spike has an initial C_p value of -0.67 and decreases to -0.87 at $8^\circ \alpha$. At $10^\circ \alpha$, the leading edge suction spike had largely diminished, and at 15° , the pressure profiles had become identical. The pressure profile mimics that of the larger porosity

configuration in rising to a second C_p peak, with a value of -0.86 at $18^\circ \alpha$. Flow visualizations displayed no appreciable differences between the other porous configuration.

For the constant rate pitching portion of the experiment, flow visualizations were taken at a freestream velocity of 20 ft/sec ($Re = 48000$). The values of α^+ used were 0.01 , 0.02 , 0.05 , 0.10 and 0.20 , ($d\alpha/dt = 0.40$, 0.80 , 2.0 , 4.0 , and 8.0 rad/sec respectively). These values were used so that direct comparisons could be made with data collected by Schreck and colleagues¹⁰. Four surface conditions were utilized: both the large and small porosity screens, and both of these screens with the paper backing in place (the non-porous condition). Selected flow visualization frames at the same α 's used for the static α conditions were again utilized for more in-depth analysis. To date, no pressure data is available for these conditions.

Fig. 8 relates DSV initiation to t^* for the four surface conditions as well as for a solid surface wing. The solid wing had a semi aspect ratio of 2 but otherwise was of a configuration similar to that used throughout the rest of this experiment. (This data is courtesy of Schreck¹⁷.) Fig. 8 indicates that the temporal variation in DSV between solid and non-porous wings was minimal for each individual α^+ . (It should be noted that DSV initiation times were determined while watching the high speed video frames which had a time between frames equal to a Δt^* of 0.16 .)

A conspicuous variation to this trend was the apparent lack of a DSV for the small open area non-porous wing configuration at $\alpha^+ = 0.01$ and 0.02 . Although flow field development proceeded in a fashion which was qualitatively similar to that described in other previous work^{12,18,19} as well as in the present study, the flow field eventually broke into one similar to those observed for the porous wing configurations as described below. For α^+ 's of 0.05 or greater, flow field development about this wing configuration behaved in a fashion similar to the solid and large open area, non-porous wing configuration.

Removal of the paper backing had a clear effect on DSV initiation phenomena. For α^+ of 0.01 and 0.02 , flow visualizations showed no clear evidence of the existence of a DSV; in its place was a turbulent layer residing over the wing surface at values of α of 0° to nearly 30° (see Fig. 10a). Qualitatively, these turbulent layers were of a similar nature for

both the large and small open area wing surfaces at each individual value of α^+ . For α^+ 's of 0.05 and greater, a distinct DSV was observed, but was initiated at significantly earlier t^* 's, as shown in Figs. 8.

A measurement of wake width for those cases in which either no DSV was found to be initiated or where wake width was not dramatically altered in the chosen frames by the presence of the DSV was made by observing the orthogonal distance between the wing's mean chord line and the first cohesive streakline above the suction surface. Fig. 11 displays these distances for α^+ of 0.01, 0.02, 0.05. These graphs show that wake widths for porous surfaces were consistently larger than for non-porous surfaces at lower angles of attack, generally at or below 20° . Beyond such α 's, wake width distributions tended to become more random due to the presence of vortical structures shed from the vicinity of the leading edge into the wake, or due to the shed DSV or related vortices.

Discussion

Fig. 3 shows that significant differences in surface pressure profiles existed between porous and non-porous surface conditions in terms of overall magnitudes, spatial extents and aft-portion attachment / separation. However, as Fig. 4 shows, pressure gradients near the leading edge remained reasonably similar for all configurations at most α 's. Such a trend indicates that flow attachment at the leading edge remains unaffected by porosity or roughness. However, the sizable difference between the non-porous and solid surface pressure profiles in Fig. 3 may indicate otherwise. Lower wing surface porosity does allow a slight peak in suction near the leading edge, but otherwise, surface pressures remained relatively unaffected by the amount of open area.

Aft of $c/4$, porosity appears to have a definite effect. Flow visualizations (Figs. 4 and 6) reveal a region of unattached flow resides over this portion of the wing for both open areas and even at $0^\circ \alpha$. Surface pressure measurements confirm this observation through the lack of any pressure recovery in this region. As α exceeded 15° , however, surface pressure profiles became nearly identical with the non-porous wing configuration while flow visualizations showed almost no differences between the two porous and non-porous wing configurations.

One possible explanation for this combination of events is that instead of an "internal circulation" similar to that described by Abbott⁹, a

flow through the internal cavity may have been created. A small mass flow through the forward portions of the porous surface and bled out the aft portions at low velocities could cause the flow to separate in this region by creating a bubble or stream of low velocity flow at the surface and thus a reasonably sized shear layer, thus accounting for the apparent turbulence. Such a flow would be created at low angles of attack, but as the exposed area of the suction surface's leading edge waned with increasing angle of attack, this internal flow would also wane and the elicited wake of the porous surface configured wing would become increasingly similar to that of the non-porous or solid surface wing. Wake width measurements (Fig. 7) appear to support this hypothesis since, at low angles of attack, wake widths are significantly larger for the porous surface configurations than for the non-porous ones.

DSV initiation times are plotted as a function of α^+ in Fig. 8. As can be seen in Fig. 8, DSV initiation for the non-porous conditions coincide with reasonable agreement with the solid wing surface condition for all but two exceptions, where no DSV is formed. It is noted that the non-dimensional time lapse between the flow visualization frames from which this data was extracted was $\Delta t^* = 0.16$, which accounts for much of the difference in initiation times for the higher α^+ values. In the cases of the non-porous low porosity surface, $\alpha^+ = 0.01$ and 0.02 , no definite DSV was formed; instead, boundary layer thickening, usually associated with the initial processes of dynamic stall, was followed by a sudden shedding of similarly structured vortical structures at a much higher frequency than associated with the DSV / tertiary vortex shedding reported in previous studies^{12,18,19}. For α^+ values at and above 0.05 , DSV initiation corresponded well with the other non-porous and solid surface cases. Thus, with the aforementioned exception, surface roughness appears to have minimal effect on DSV initiation.

Adding the effect of porosity caused a number of differing effects on DSV initiation, all of which were highly dependent on α^+ . For values of α^+ of 0.01 and 0.02 , no DSV was initiated. In its place was a series of vortical structures which emanated from the region of the leading edge. These structures were of approximately the same nature at each individual α^+ for both porous wing configurations. Hence, it becomes apparent that the addition of the open surface caused a subsequent lack of vorticity

accumulation in the vicinity of the leading edge at these lower pitch rates, possibly due to attenuated suction peaks at the leading edge and / or the impediment of reverse flow in the boundary layer through the addition of the porous surface.

As α^+ increased to 0.05 and beyond, an increasingly apparent DSV was noted. In these cases, as can be observed in Fig. 8, the DSV was initiated significantly earlier in the pitch motion compared to the non-porous and solid wing configurations. The computational work of Visbal²⁰ indicates that the presence of a disturbance within the boundary layer near the leading edge, whether originating at the trailing edge and transported forward through reverse flow or originating at the leading edge, is the apparent trigger of DSV initiation. Thus, a logical conclusion is that the flow field elicited by the porous surface wing either is much more sensitive to such disturbances or casts such disturbances earlier in the pitch motion than for the non-porous or solid surface wings into the vicinity of accumulating vorticity.

Conclusions

A study of the effects of the addition of partially open surface screens to a NACA 0015 wing section resulted in the partial characterization of the flow fields elicited by such a wing at static angle of attacks and during dynamic constant rate pitch motions.

At static angles of attack, the wing with porous suction surfaces showed a significant decrease in surface pressure values at angles of attack below 15° - 20°, and experienced aft portion separation. When the porous surface was backed, creating a non-porous surface, surface pressure values increased over those of the porous surface, and the flow remained fully attached; however, pressure gradients ahead of c/4 remained of the same magnitude in all cases. Once angles of attack exceeded 20°, little differences existed between flow fields elicited by porous and non-porous configurations. A possible explanation for these observations is a mass flow entering the underlying cavity in the vicinity of the leading edge and bleeding out aft of c/4. The non-porous surface does experience a significantly greater C_p magnitude at 0° α than the theoretical solid surface does, indicating a possible roughness effect or procedural error.

During dynamic pitch motions, dynamic stall vortex initiation was induced earlier in the pitch motion for α^+ values at or greater than 0.05,

while no DSV was apparent at test α^+ values below 0.05. Open area percentages had little noticeable effect on forced flow field development at each individual pitch rate. Placing a non-porous backing on the porous surface caused the elicited DSV structures to closely mimic those elicited by solid surface wings.

The addition of a perforated metal sheet to a wing appeared to cause the most flow field alteration aft of $c/4$, most notably at low angles of attack, for both static and dynamic conditions. Low velocity mass injection in this region apparently caused the flow to be separated at $0^\circ \alpha$, apparently causing the host of variations in flow field developments observed in this study. Ahead of $c/4$, and most notably in the first 10% c , static α surface pressure measurements indicate that, although a shift in stagnation point is observed, pressure gradients were maintained in magnitude, indicating that overriding flow forcing functions retain at least some of their importance in the overall flow descriptions.

Acknowledgments

This work was sponsored by the USAF Office of Scientific Research under the Graduate Student Research Program, administered by Research and Development Laboratories of Culver City, CA. The technical support of S. Schreck, M. Luttges, B. Hatfield, D. Bunker, J. Lovato, J. Dougherty, J. Klinge and M. Robinson as well as the remainder of the USAFA/DFANL and FJSRL staff was greatly appreciated.

Bibliography

1. Gad-el-Hak, M, and Blackwelder, R.F., "Selective Suction for Controlling Bursting Events in a Boundary Layer", AIAA J., Vol. 27, March 1989, pp 308-14.
2. Simpson, R.L., Moffat, R.J., and Kays, W.M., "The Turbulent Boundary Layer on a Porous Plate: Experimental Skin Friction with Variable Injection and Suction", Int'l. J. of Heat and Mass Transfer, Vol. 12, July 1969, pp 771-89.
3. Taylor, J.W.R. (ed.), Jane's All the World's Aircraft, 1966-67, McGraw Hill, New York, 1966, pp. 299-300.
4. Holt, D.J. (ed.), "Laminar Flow Testing Examined", Aerospace Engineering, Vol. 9, Jan. 1989, Society of Automotive Engineers, pp. 13-16.
5. Chen, C.L., Chow, C.Y., Van Dalsem, W.R., and Holst, T.L., "Computation of Viscous Transonic Flow over Porous Airfoils", J. of Aircraft, Vol. 26, December, 1989, pp. 1067 - 75.

6. Savu, G. and Trifu, O., "Porous Airfoils in Transonic Flow", AIAA J., Vol. 22, July 1984, pp 989-91.
7. Krogmann, P., Stanewsky, E., and Thiède, P., "Effects of Suction on Shock/Boundary-Layer Interaction and Shock-Induced Separation", J. of Aircraft, Vol. 22, January 1985, pp 37-42.
8. Soderman, P.T., Flow-Induced Resonance of Screen-Covered Cavities, NASA Technical Paper 3052, October, 1990.
9. Abbott, I.H., "Experiments with an Airfoil Model on which the Boundary Layer Is Controlled without the Use of Supplementary Equipment", NACA TN 371, Langley Memorial Aeronautical Laboratory, April, 1931.
10. Schreck, S.J., Addington, G.A., and Luttges, M.W., "Flow Field Structure and Development Near the Root of a Straight Wing Pitching at Constant Rate", AIAA Paper 91-1793, June 1991.
11. Robinson, M., Walker, J., and Wissler, J., "Unsteady Pressure Measurements on a Pitching Rectangular Wing", Proceedings of Workshop II on Unsteady Separated Flows, USAFA, CO, 1988, pp 225-37.
12. Helin, H.E., Experimental Studies on the Dynamic Development and Control of Unsteady Separated Flows, Doctoral Thesis, University of Colorado, 1986.
13. Abbott, I.H., and Von Doenhoff, A.E., Theory of Wing Sections, Dover Press, New York, 1959, p 324.
14. Jacobs, E.N., and Sherman, A., "Airfoil Section Characteristics as Affected By Variations of the Reynolds Number", NACA Report No. 586, 23rd Annual NACA Reports, 1937, pp 577-611.
15. Batchelor, G.K., An Introduction to Fluid Dynamics, Cambridge University Press, 1967, pp 437-41.
16. Milne-Thompson, L.M., Theoretical Aerodynamics, 4th ed. Dover Publications, 1958, pp 106-08.
17. Schreck, S., Private Communications, 25 May - 06 August 1991.
18. Helin, H.E., Robinson, M.C., and Luttges, M.W., "Visualization of Dynamic Stall Controlled by Large Amplitude Interrupted Pitching Motions", AIAA Paper No. 86-2282-CP, August 1986.
19. Robinson, M., Helin, H.E., Gilliam, F., Russel, J., and Walker, J., "Visualization of Three-Dimensional Forced Unsteady Separated Flow", AIAA Paper No. 86-1066, May 1986.
20. Visbal, M., "On the Formation and Control of the Dynamic Stall Vortex on a Pitching Airfoil", AIAA Paper No. 91-0006, January 1991.

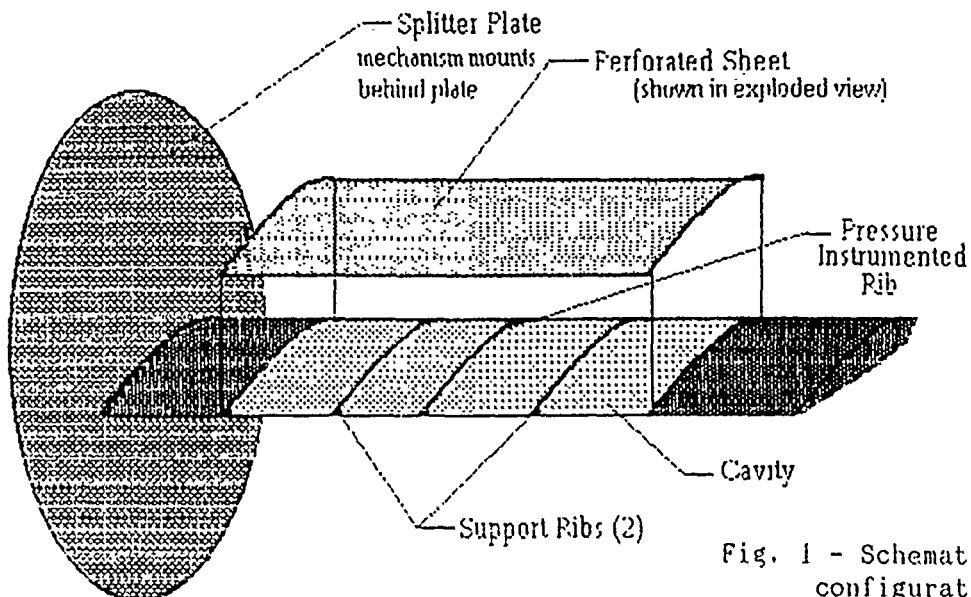


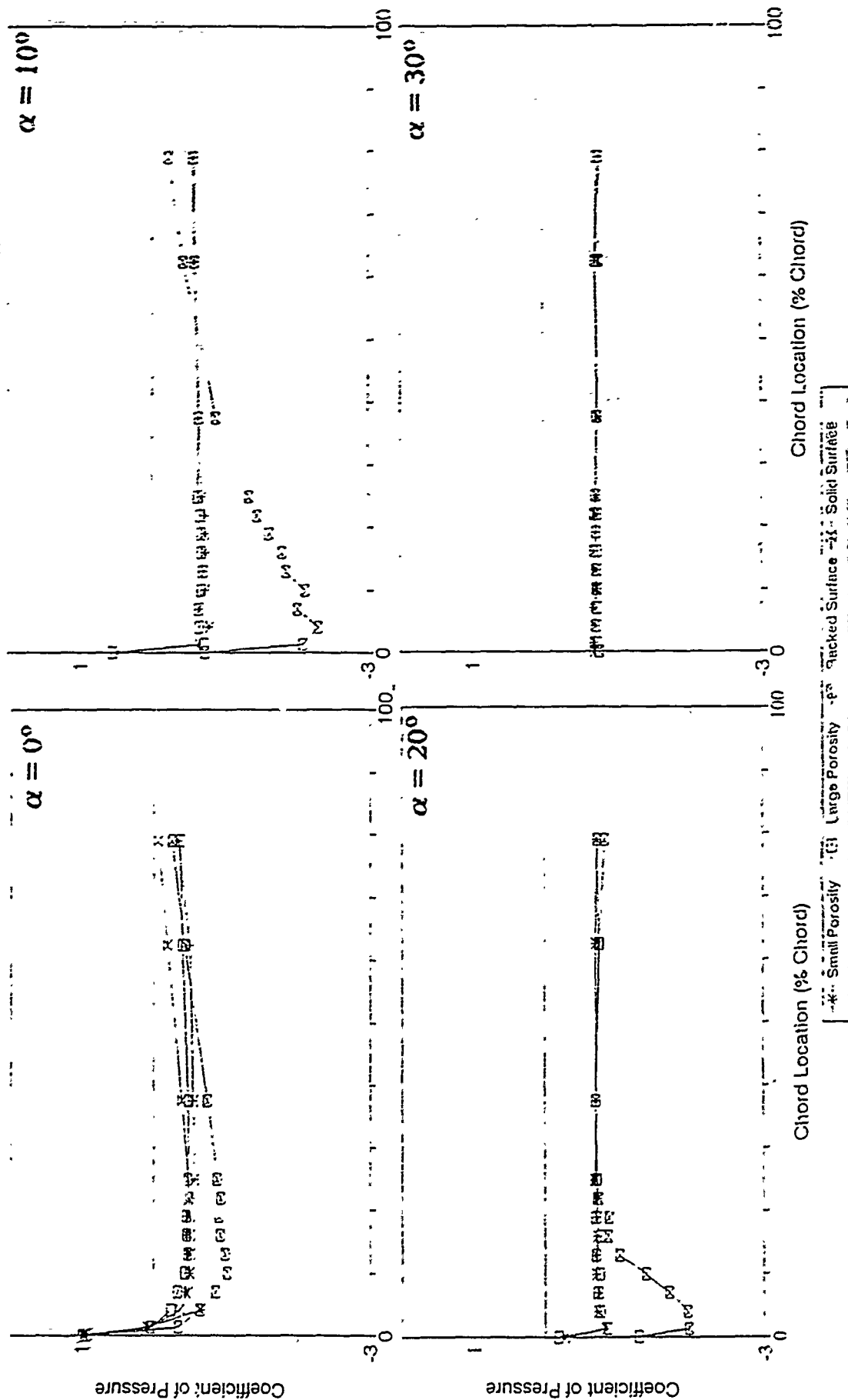
Fig. 1 - Schematic of porous wing configuration. Not to scale.

| Sheet Type | Hole Diameter | Center-to-Center Spacing | Percent Open Area |
|-----------------|---------------|--------------------------|-------------------|
| Small Open Area | 0.016 in | 0.026 in | 30 |
| Large Open Area | 0.038 in | 0.050 in | 45 |

Table 1 - Open area and hole size data for the porous surface sheets.



Fig. 2 - Placement of pressure taps along NACA 0015 cross section.



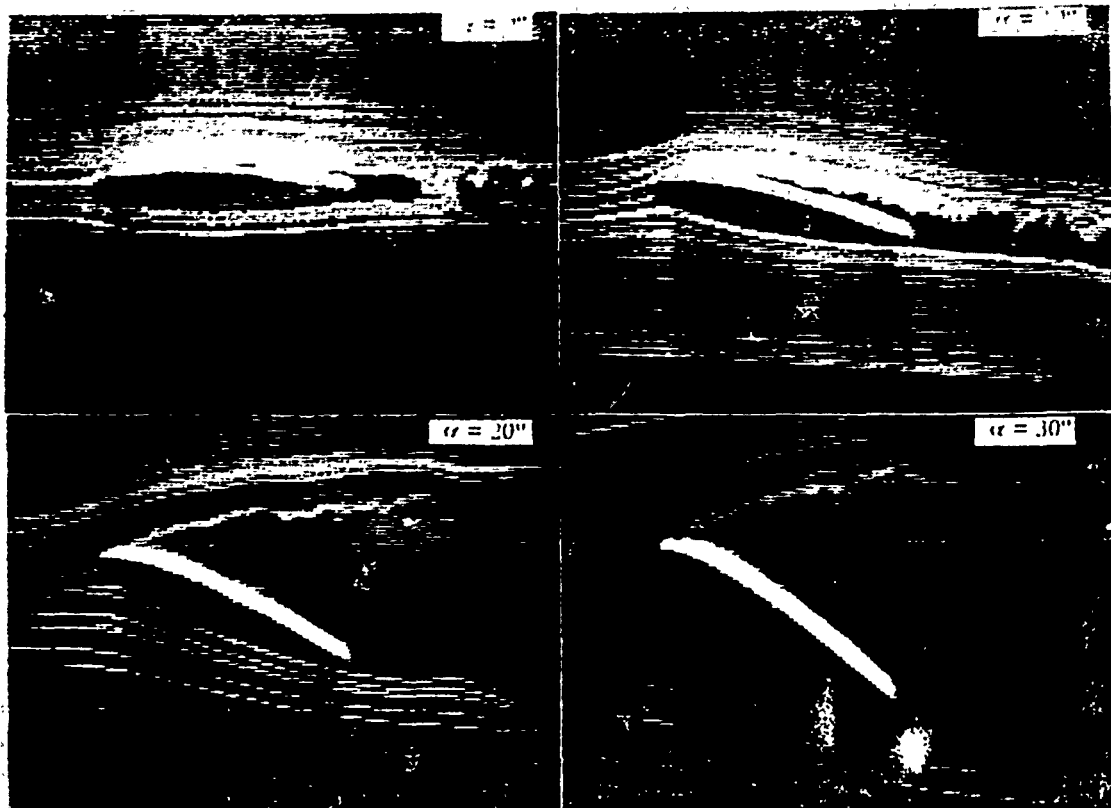


Fig. 1 - Flow visualization for the porous, large open area sheet wing configuration at the angles of attack shown.

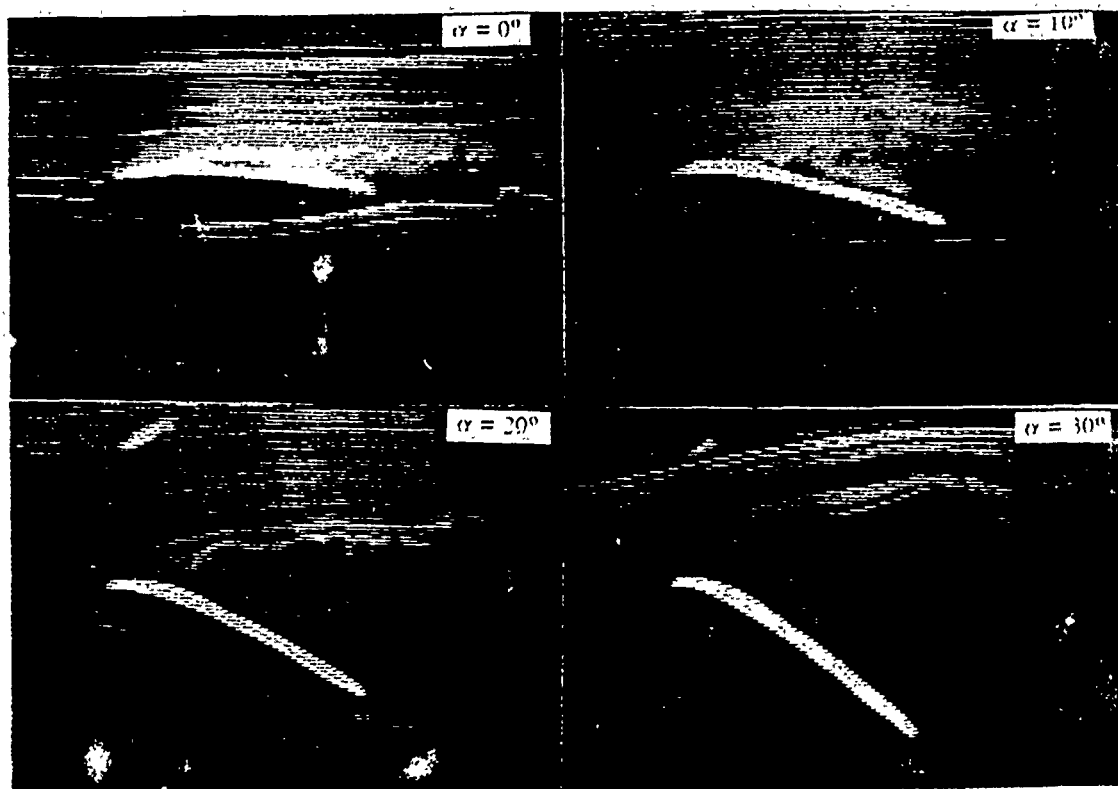


Fig. 2 - Flow visualization for the porous, large open area sheet wing configuration at the angles of attack shown.

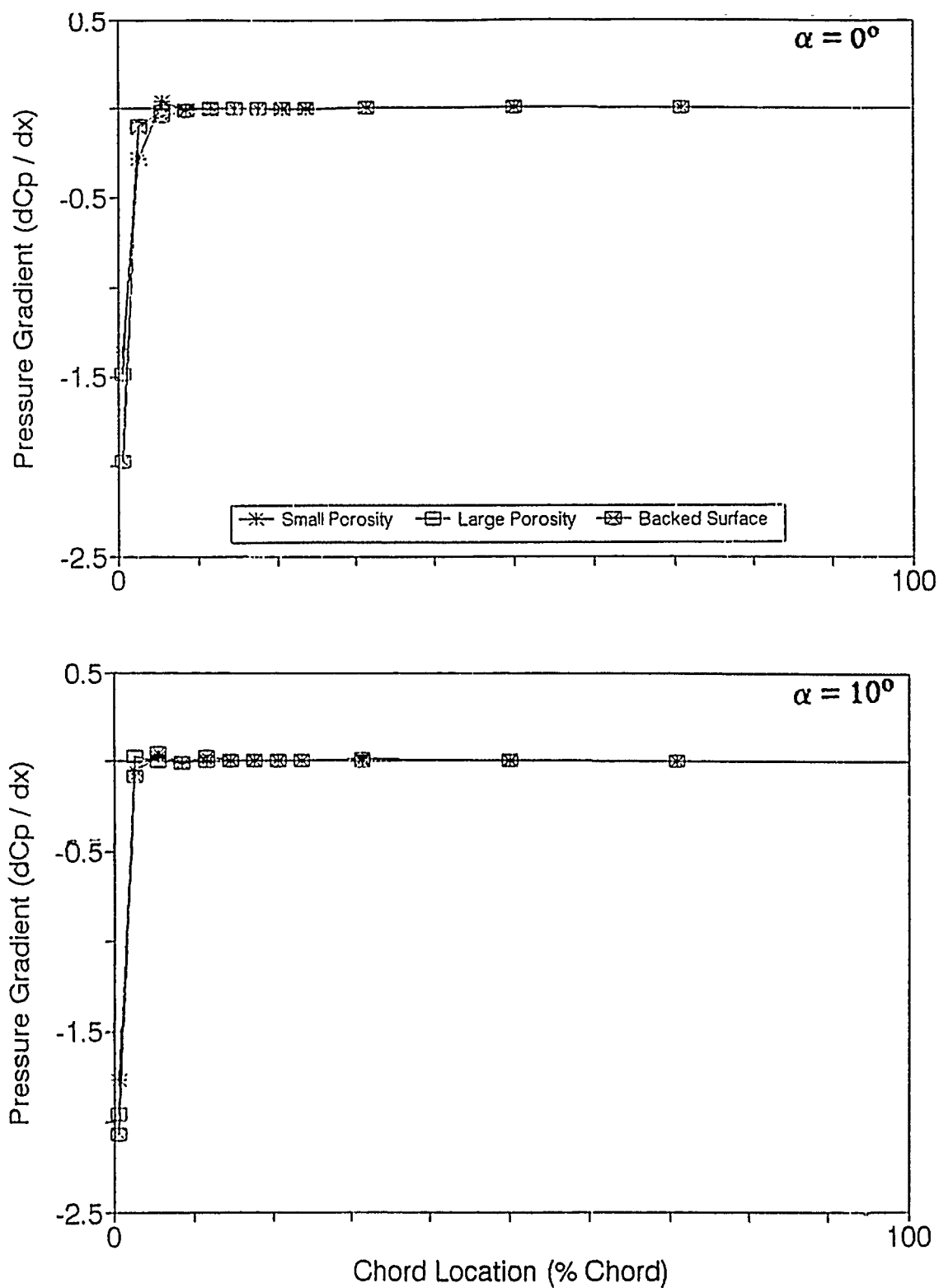


Fig. 5 - Plots of dC_p/dx taken from the values plotted as Fig. 3.
Gradient values are plotted at the interval mean chord station.

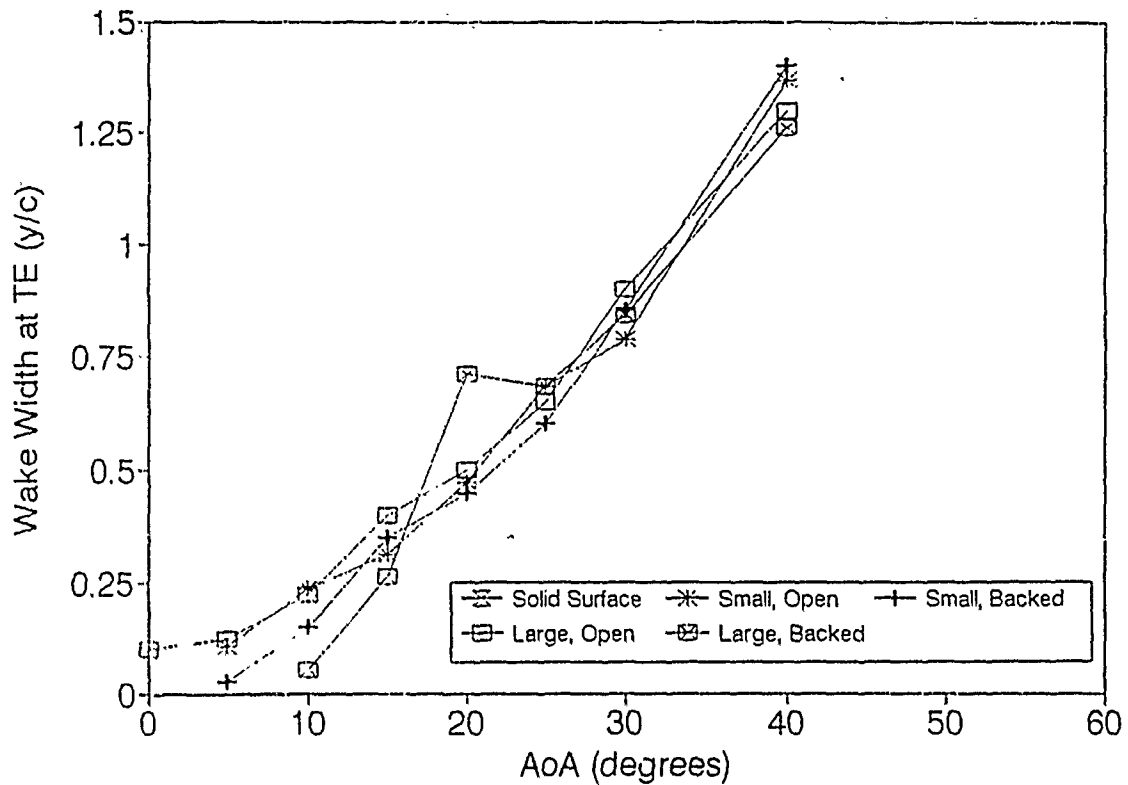


Fig. 7 - Wake width at the trailing edge for the four surface configurations at static angles of attack.

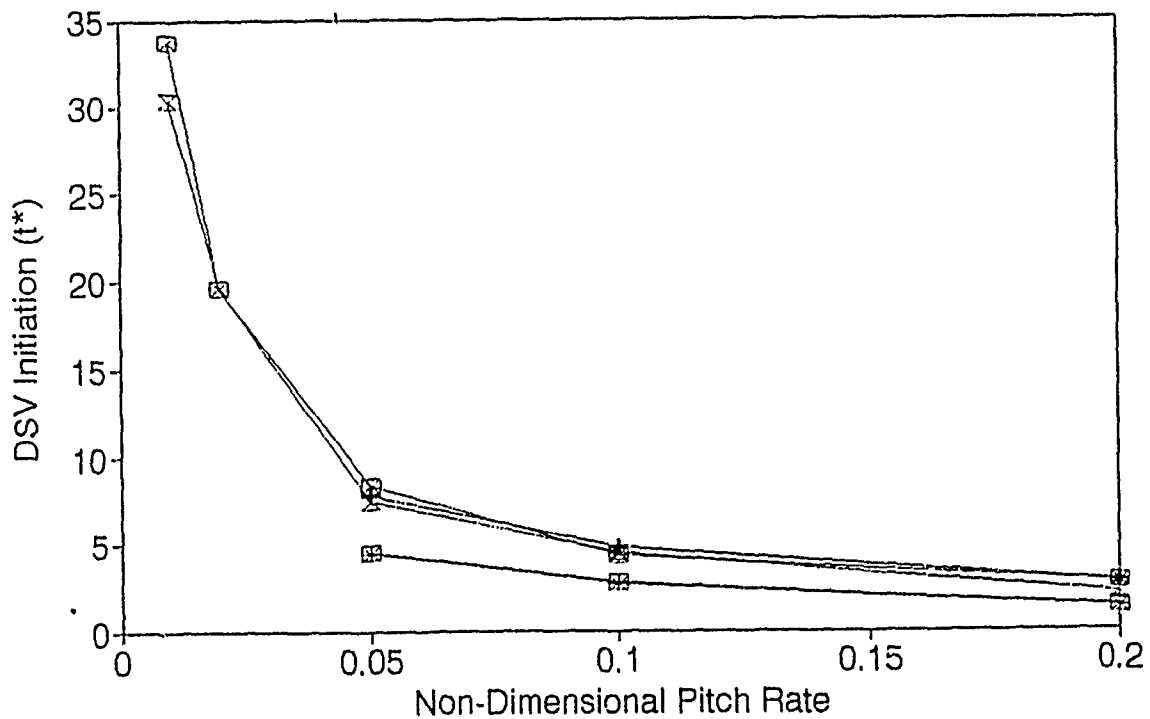


Fig. 8 - DSV initiation t^* for 5 surface configurations during constant rate pitch motions.

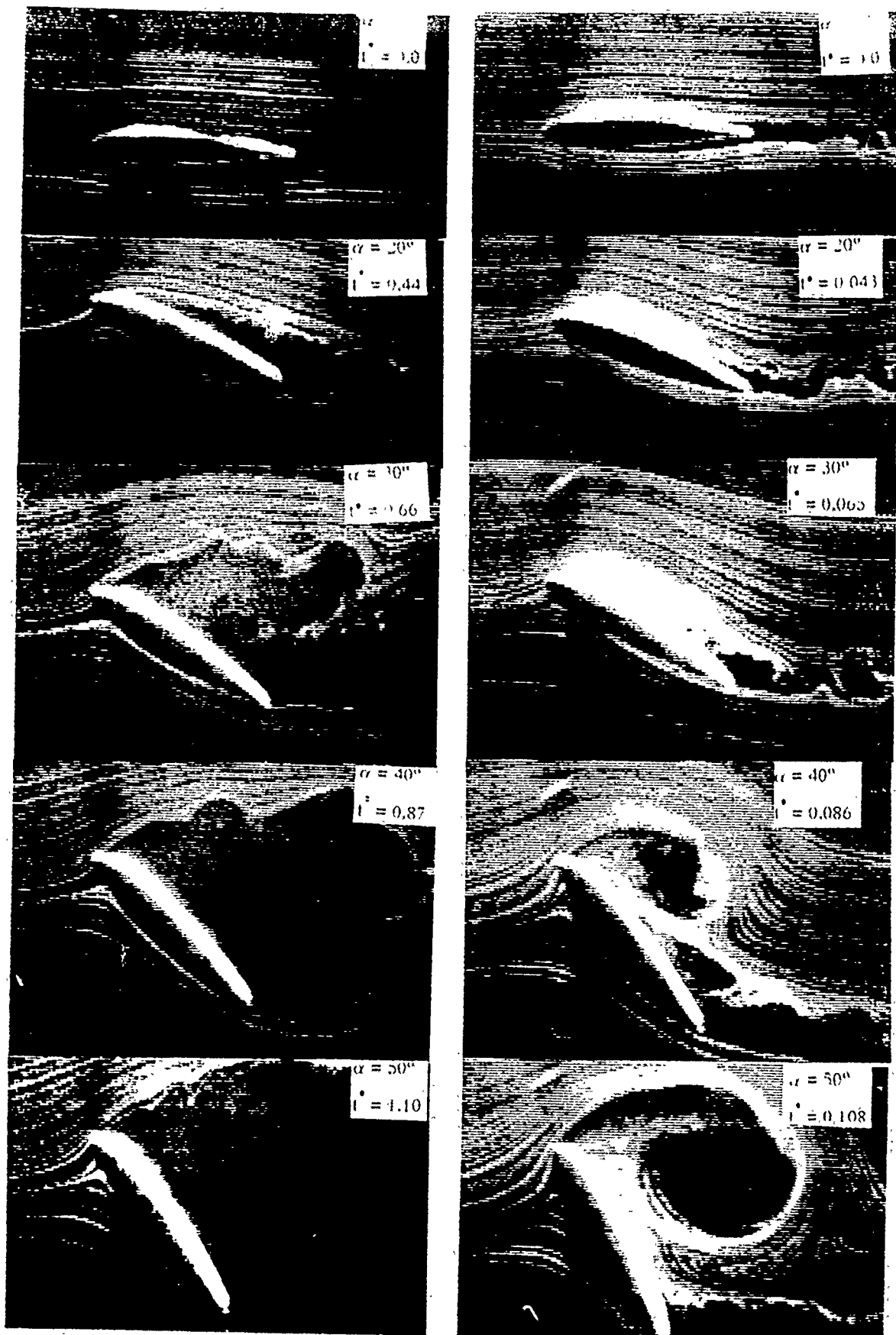


Fig. 9 - Flow visualization for the non-porous, large open area sheet
 ring concentration at $\alpha = 10^\circ$ (left) and 50° (right) at the Re
 values indicated in the figure.

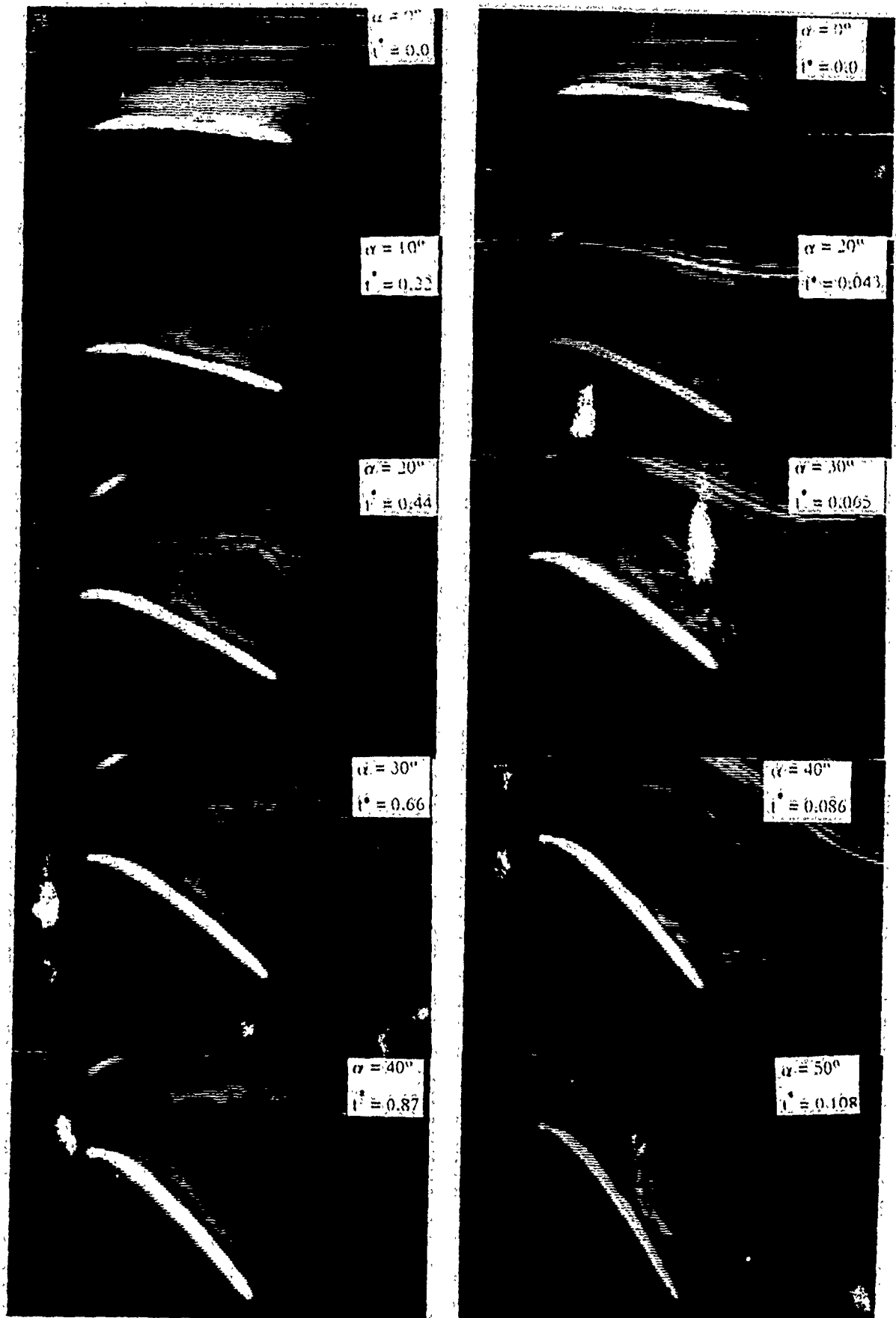


Fig. 10 - Flow visualization for the porous, large open area sheet wing configuration at $\alpha^* = 0.02$ (left) and 0.20 (right) at the α and t^* shown.

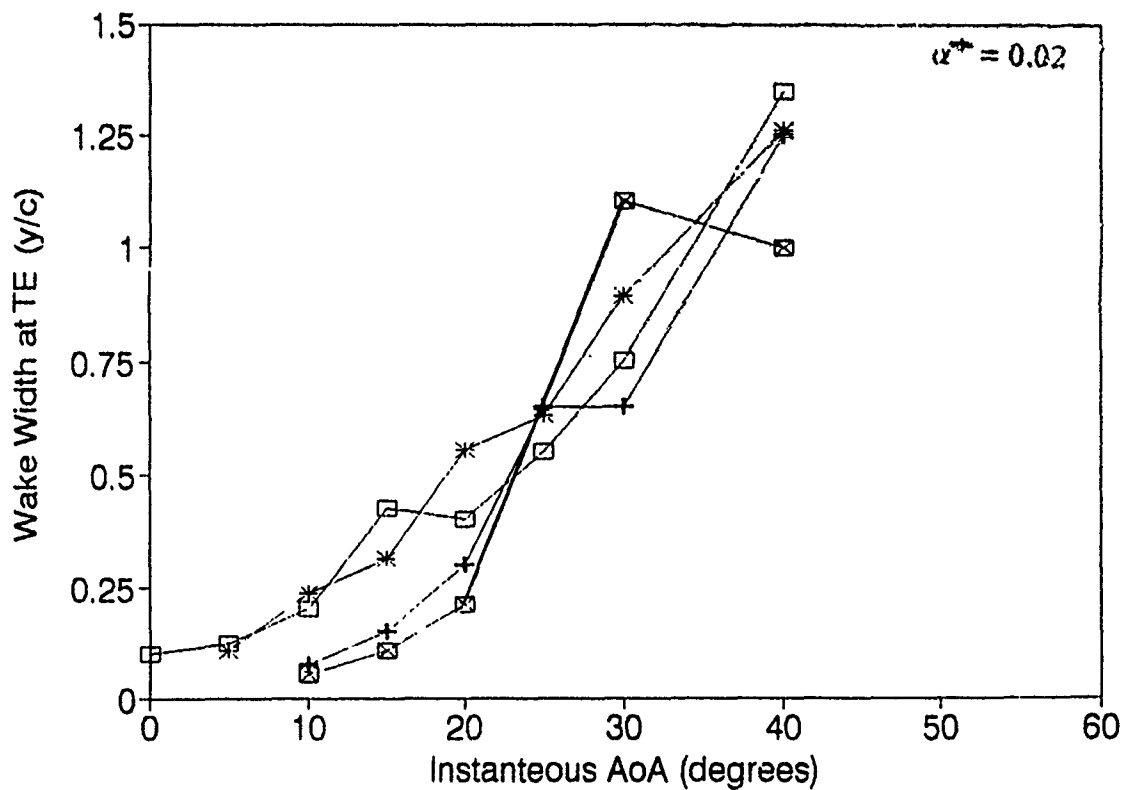
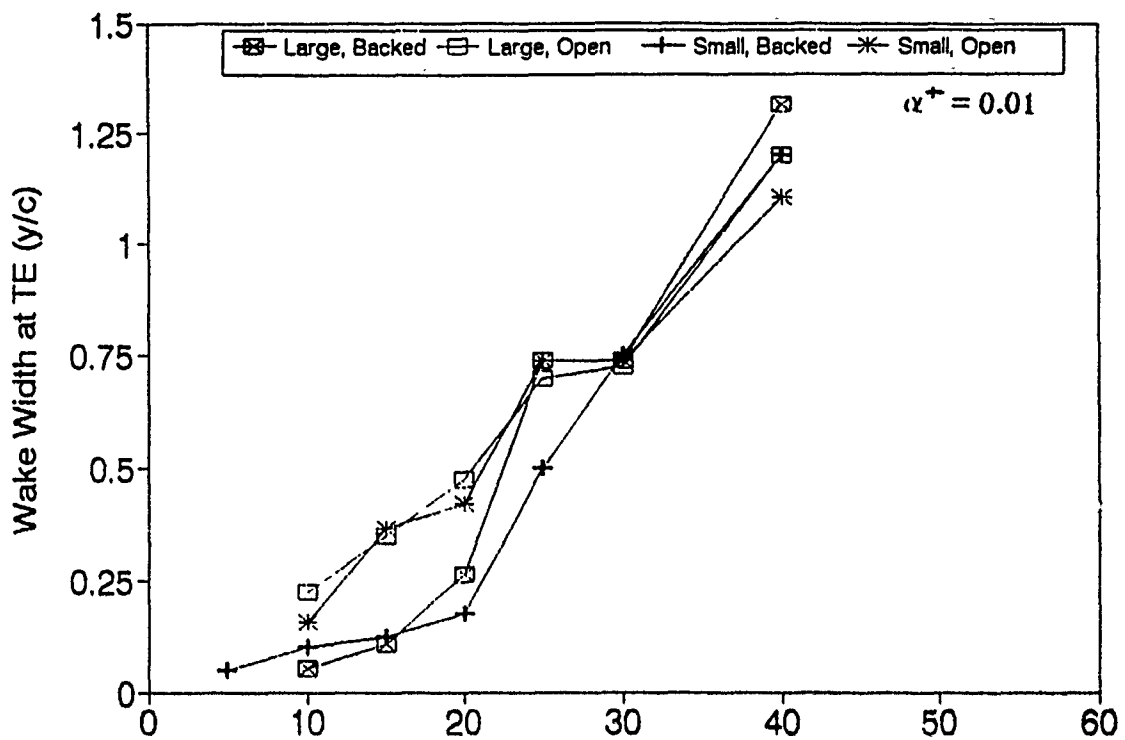


Fig. 11 - Trailing edge wake widths for 4 surface configurations at the indicated non-dimensional pitch rates.

ACTIVE/PASSIVE CONTROL SYNERGISM FOR A PLANAR TWO-DIMENSIONAL TRUSS

Jeffrey D. Curtis
Graduate Research Fellow
State University of New York at Buffalo

Abstract

An experimental example of active/passive control synergism was accomplished for a planar truss system using visco-elastic struts and air-jet thrusters. It was shown that the passive dampers can eliminate a third mode instability and increase the structural damping with only a relatively small loss in stiffness.

Introduction

Future large space structures will be constructed of light-weight members. These structures will tend to vibrate easily, and therefore must have damping mechanisms built-in to quickly and effectively eliminate unwanted vibrations. Various techniques can be used to damp these vibrations. One approach is to use a combination of active control and passive damping to suppress vibrations in a wider frequency range than either method could achieve by itself.

To better understand the dynamic interactions between flexible structural systems, air-jet thrusters, and visco-elastic members, the control of a flexible planar truss was studied. The setup included a 20 bay planar truss, 7.07 meters long, supported on ball bearings to restrict motion to the horizontal plane. In this study, the active control comes from two pairs of air-jet thrusters mounted at the midpoint and the free end, while the passive damping results from a visco-elastic member appropriately placed in the structure. The problem with the thrusters results from the time delay in the response of the

servovalve. These time delays in the response time of the servovalves induced marginal stability at or near the third mode, hence, the thrusters cannot effectively damp structural resonances at frequencies above 60 Hz.

Discussion

The 20 bay planar truss, shown in Figure 1, consisted of Meroform aluminum struts and steel nodes, with steel bars across every bay (vertically, in the diagram) to add weight and keep the truss supported on ball bearings. The air-jet thrusters consisted of a solenoid valves made by Mac Valves, Inc., model 113B551BAAA, fitted with a nozzle perpendicular to the truss. Air was supplied to the thrusters at 64 psi from the building air supply.

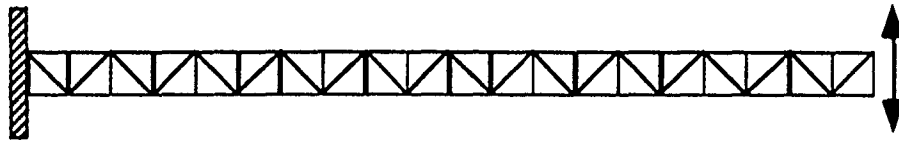


Figure 1. 20 bay planar truss.

The marginal stability of the thrusters at the third mode was explained by Hallauer and Lamberson¹ (1989). They reasoned that a $\pi/2$ phase lag in actuator could cause control degradation or instability near the frequency of interest. This is equivalent to a frequency of $\frac{1}{4\tau}$, where τ is the time delay in the servovalve of the air-jet thrusters. A thruster servovalve which had a 14 ms time delay, hence corresponding to a frequency of 17.9 Hz. The frequency response function for the thruster gave a peak at 16.1 Hz. The actual thrusters used on the planar truss were shown to have a 10 ms delay, corresponding to 25 Hz. The third mode for the truss is 24.6 Hz.

One way to eliminate the thruster problem is to use a visco-elastic member to increase third mode damping. This member will be tested in two ways. First, it will be attached in parallel with a normal strut in an attempt to add damping without significant loss in stiffness (see Figure 2). The second test involves replacing an original strut with a passively damped strut, both alone and with a regular strut in parallel (see Figures 3 and 4). This should provide greater damping than the previous method, but could also reduce stiffness.

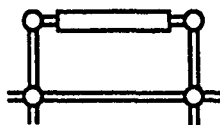


Figure 2.

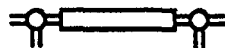


Figure 3.

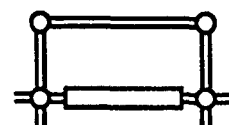
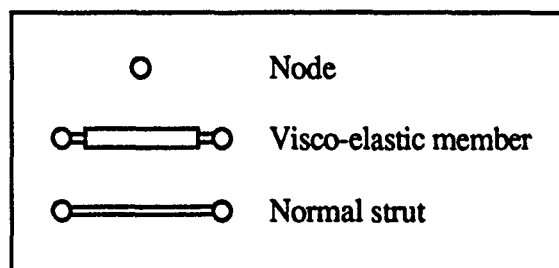


Figure 4.



The overall goal is to add enough damping to the third mode to eliminate the thruster problem without large changes in natural frequency or stiffness. This should be accomplished by placing the visco-elastic strut in a location that affects mode 3 more than the other modes. The optimal locations for the strut are found from the strain energy data given for each mode from I-DEAS. Figure 5 shows the best placements for the first four modes: the numbers represent the mode and the arrows indicate the best locations for the third mode.

To determine the effect of the visco-elastic member on the truss, a frequency response function (FRF) will be performed for each configuration using a random noise

input on a Tektronix 2630. Free decay damping calculations will also be done to check for changes in damping. These results will be compared to a "baseline" FRF after averaging the results of a few trials.

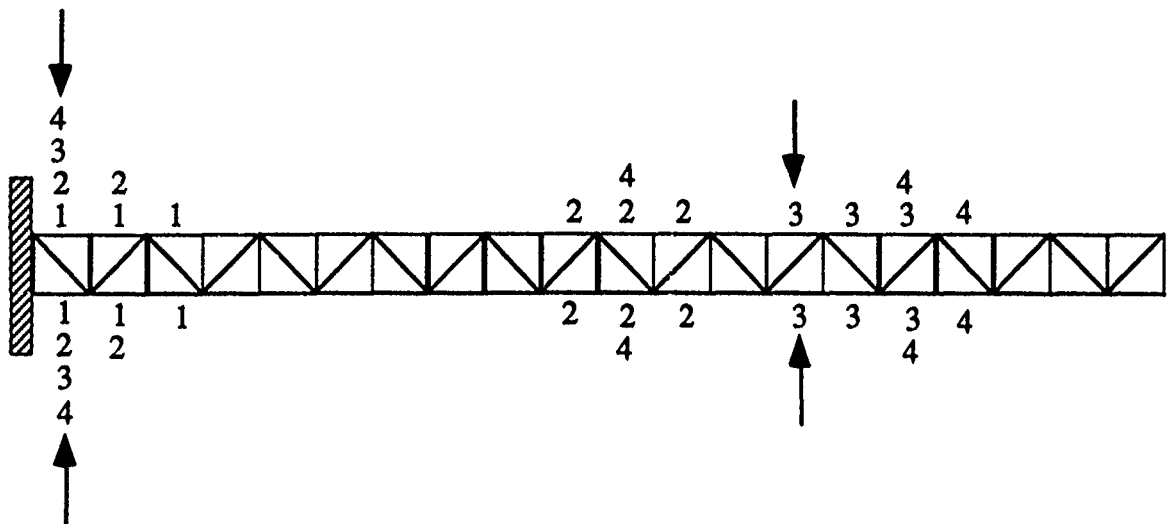


Figure 5. Optimal locations for visco-elastic strut placement.

Results

Two locations opposite each other, about two-thirds down the length of the truss from the base, were selected for a comparative study (see Figure 6). This particular bay was selected over the first bay (at the base) because placing a passively damped strut at that location reduced the stiffness in every mode beyond an acceptable limit.

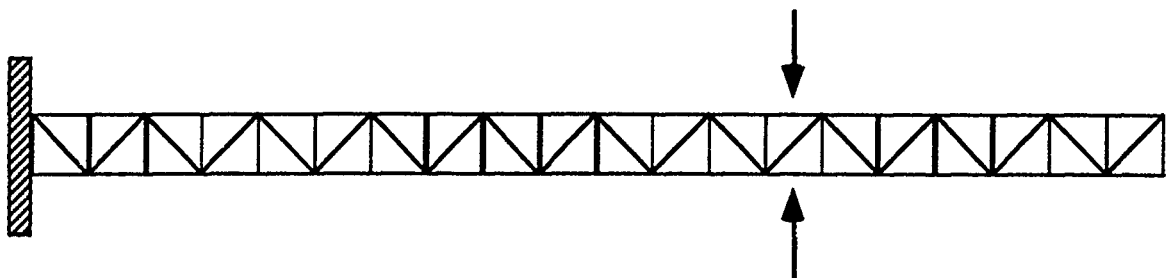


Figure 6. Locations where visco-elastic member was placed.

Four aspects of the truss were studied: the change in natural frequency, peak amplitude, stiffness, and damping. The natural frequency and peak amplitude (converted from decibels to actual values) were taken from the FRF and compared to the baseline trial. The change in stiffness was calculated from the following equation using the change in frequency:

$$\% \text{ change in } K = (1 + \% \text{ change in } \omega)^2 - 1 \quad (1)$$

The damping was calculated from free decay graphs after excitation at each mode. The method of logarithmic decrement was used to get the value of ζ . Assuming viscous damping, ζ is calculated as follows:

$$\zeta = \frac{\delta}{\sqrt{\delta^2 + 4\pi^2}} \quad (2)$$

where $\delta = \frac{1}{n} \ln \left(\frac{x_0}{x_n} \right)$ and n represents the number of cycles between the peaks x_0 and x_n on the graph.

Placing a passively damped member in a parallel load path (see Figure 2) did not have any significant effect on the structure, and therefore no further experiments were performed in that configuration. However, replacing a normal strut with a passive member had a significant effect on the structure, particularly with respect to damping. Three configurations were used to test the effects of the visco-elastic strut. First, just a passive strut was inserted (see Figure 3), then a parallel load consisting of a normal strut was added for both a "short" and "long" separation from the truss (see Figure 7). These configurations were used at both of the locations shown in Figure 6. It was found that there was no significant difference between configurations or locations. The only effect of

adding a normal member in parallel to a visco-elastic member was the slight increase in first mode frequency, stiffness, and peak amplitude. However, no other modes or aspects displayed any similar trends. A summary of the results appears in Table 1 for the FRF taken from the tip of the truss.



Figure 7. Normal member in parallel with a visco-elastic member.

Table 1. Results of comparative study.

| | Mode 1 | Mode 2 | Mode 3 | Mode 4 |
|---------------------------|-----------------------------------|----------------------------|----------------------------|----------------------------|
| Frequency | 2.6% Decrease to 3.2% Increase | 2.2% to 3.7% Decrease | 2.1% to 3.8% Decrease | 0.1% to 0.9% Decrease |
| Peak Amplitude | 14.5% to 21.3% Increase | 41.4% to 48.0% Decrease | 42.7% to 53.3% Decrease | 18.3% to 36.0% Decrease |
| Stiffness | 5.2% Decrease to 6.3% Increase | 4.4% to 7.2% Decrease | 4.2% to 7.4% Decrease | 0.3% to 1.8% Decrease |
| Damping | 1.8% to 13.5% Increase | 49.0% to 53.4% Increase | 60.0% to 63.7% Increase | 4.8% to 17.1% Increase* |

* Indicates possible inaccurate data

The frequency change with the passive damper was less than $\pm 4\%$ for all modes, resulting in a change in stiffness of less than $\pm 8\%$. The relatively small change was necessary to preserve the analytical model. The amplitude of the first mode, however, showed an increase, but the other modes exhibited a much larger decrease. The second and

third modes, in particular, had very large decreases. A similar trend occurred in the modal damping. The change in first mode damping, as expected, was very small, but modes 2 and 3 had large increases. The damping values for the fourth mode were not very reliable because the amplitude of vibration approached the magnitude of the noise in the system. The percentages given for these values are only approximate, but they also showed an increase.

One very interesting result was that the changes in damping and amplitude of mode 2 was almost the same as mode 3. Referring to Figure 5, some optimal locations for second mode damping occur two bays from good locations for the third mode. Thus, inserting a visco-elastic member at one of the positions indicated in Figure 6 was not only effective at increasing the damping in mode 3, but also had a significant beneficial effect on mode 2.

The result of the addition of the passively damped strut was the elimination of the marginally stable thruster problem. Without the visco-elastic member, closed-loop control of the truss at the third mode resulted in the response shown in Figure 8. Rather than damping the vibration, the air-jet thrusters continued to excite the structure at the third mode. In this situation, the system was marginally stable. When the passive damper is inserted, the added damping prevents the thrusters from exciting the truss when they should be controlling it (see Figure 9). Note that before the control is turned on, the third mode amplitude of vibration is reduced by almost 80 percent in the structure with the passive damper. This reduction in amplitude, along with the added damping, was enough to overcome the third mode control problem.

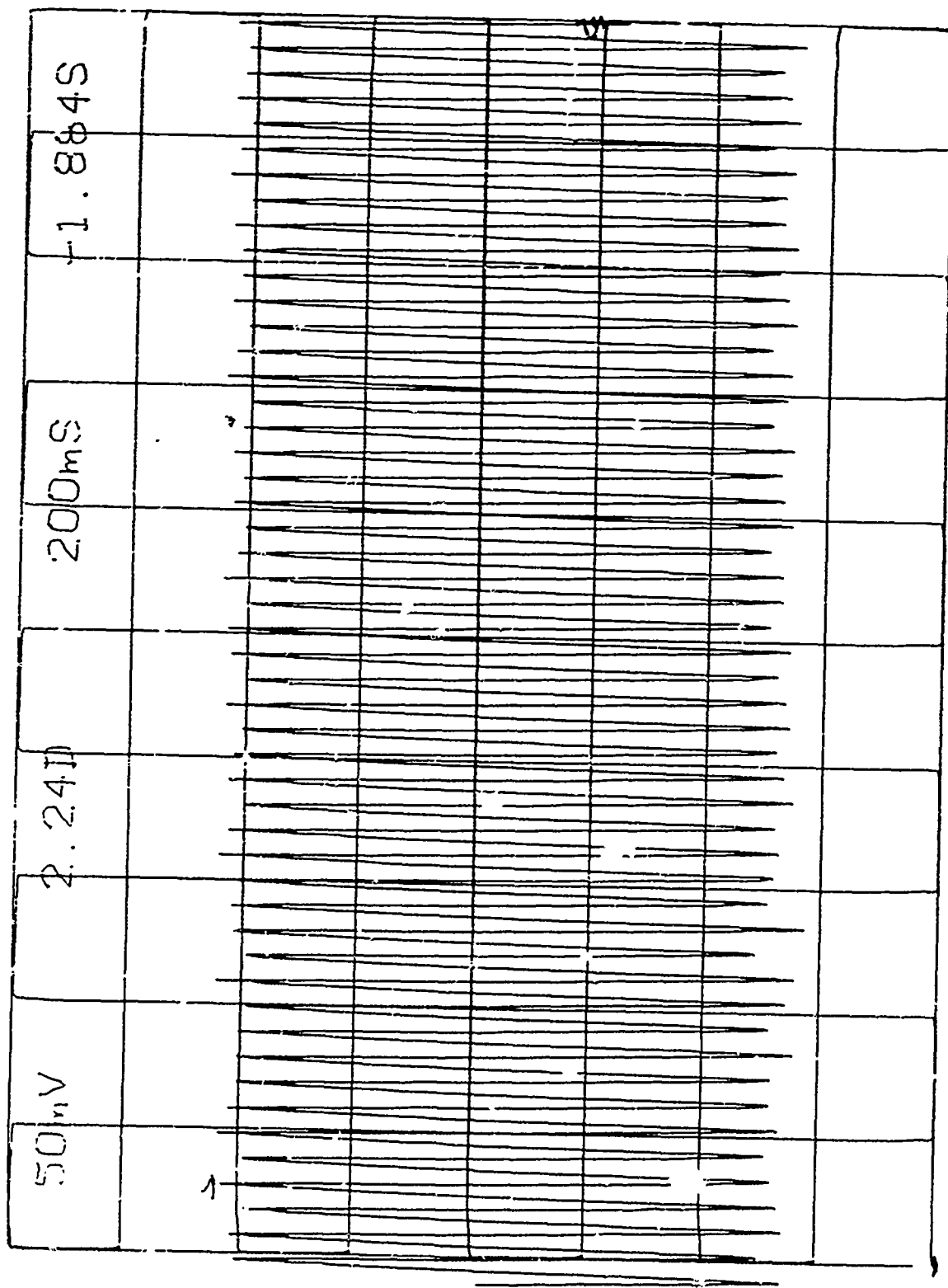


Figure 8. Closed-loop response without the visco-elastic member at the third mode.

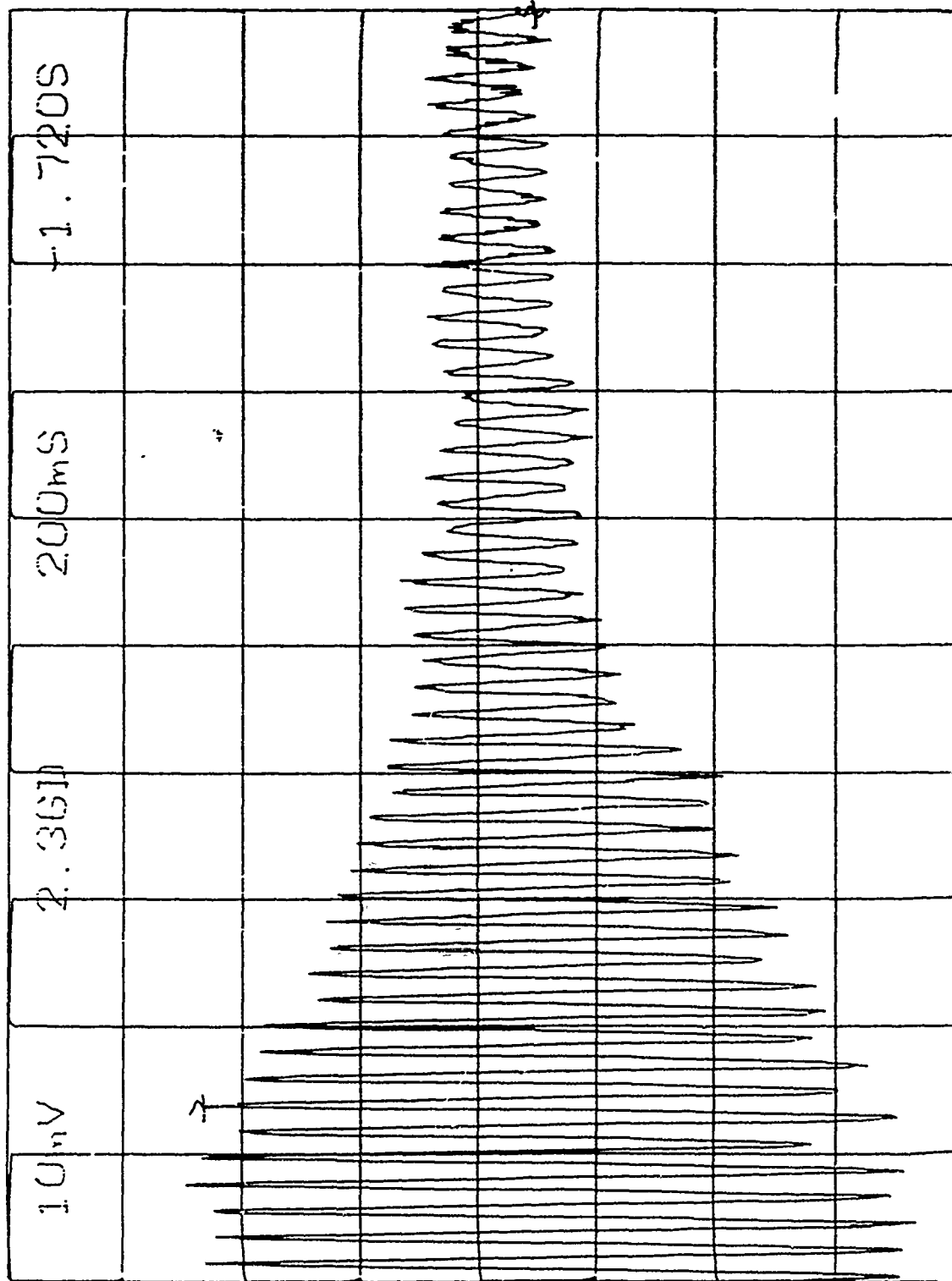


Figure 9. Closed-loop response with the visco-elastic member at the third mode.

Conclusions

The active control problem in the third mode of the 20 bay planar truss was attributed to the time delay in the servovalves of the air-jet thrusters. It was shown that a time delay of τ seconds could cause control instabilities at or near the frequency $\frac{1}{4\tau}$, and in the case of the truss this occurred at the third mode (about 25 Hz).

The visco-elastic member performed as expected, eliminating the thruster problem without a significant change in the natural frequency or stiffness of the first four modes. In addition to an increase in third mode damping, there was also a large increase in the second mode. The only drawback of inserting the damper was the increase in the peak amplitude of the first mode FRF. However, the benefits of the visco-elastic member outweigh this increase.

Other solutions to this problem include replacing the current servovalves with valves that have a time delay that will not cause control difficulties near a structural mode. Another solution is to use a new control law, perhaps with a different compensator, that will prevent this type of problem.

Acknowledgements

This research was conducted at the Frank J. Seiler Laboratory located at the United States Air Force Academy in Colorado Springs, Colorado. This work was sponsored by the Air Force Office of Scientific Research.

References

1. Hallauer, W.L. and Lamberson, S.E., "Experimental Active Vibration Damping of a Plane Truss Using Hybrid Actuation," AIAA Paper No. 89-1169, 30th Structures, Structural Dynamics, and Materials Conference, Mobile, Alabama, April 3-5, 1989.

OPTIMAL PLACEMENT OF PASSIVE DAMPERS VIA SIMULATED ANNEALING

Tami A. Hamernik
Graduate Research Assistant
State University of New York at Buffalo

Abstract

The study of placing a finite number of viscoelastic members in a large structure to provide damping for selected modes of vibration was examined. When the locations available in the structure for placement of these passive members are spatially discrete, the problem becomes a combinatorial optimization, much like the classic traveling salesman problem. For this case study, a simulated annealing algorithm was used where the cost function was the energy dissipation rate generated by a random placement of the viscoelastic members. Near optimal solutions were found and presented for the third mode of vibration in a twenty-bay planar truss structure.

Introduction

Many large space structures will tend to be lightweight and thus prone to vibrations. Active control can be used on these structures to damp unwanted vibrations, but active controllers can also lead to instabilities in the structures. For example, forced vibration tests were performed on a twenty-bay planar truss structure (80 members) to examine the structural dynamic characteristics of the truss. When the third mode was examined, the use of active control made this mode marginally stable. Passive dampers were introduced into the structure in order to stabilize the third mode.

To maximize the effect of the passive members, an optimization scheme was created based on the placement of the passive members in the structure. When the available locations for placement of the passive members are spatially continuous, a gradient-based

optimization method is sufficient. However, when these locations are spatially discrete, the optimization becomes a combinatorial problem. Since the evaluation of every combination for a finite number of passive member placements was not feasible, a heuristic-based technique was developed that renders near optimal solutions. A simulated annealing approach was used in this study to handle the selection of passive member placement.

Discussion

The equation of motion of the n -DOF system with viscous damping can be shown in matrix form as

$$M\ddot{\mathbf{x}} + D\dot{\mathbf{x}} + K\mathbf{x} = 0 \quad (1)$$

The modal decomposition of equation (1) yields the form

$$I\ddot{\mathbf{q}} + \text{diag}\{2\zeta_i\omega_i\}\dot{\mathbf{q}} + \text{diag}\{\omega_i^2\}\mathbf{q} = 0 \quad (2)$$

which is the closed-loop system equation. The state space representation of equation (2) becomes

$$\dot{\mathbf{z}} = A\mathbf{z}$$

where $\mathbf{z}=[\mathbf{q} \ \dot{\mathbf{q}}]^T$ and

$$A = \begin{bmatrix} 0 & I \\ -\text{diag}\{\omega_i^2\} & -\text{diag}\{2\zeta_i\omega_i\} \end{bmatrix}$$

This allows the effect of passive member damping to be easily inserted into the optimization algorithm, where the energy dissipation rate is used as the performance criterion. Using equation (2), the total energy stored within the system is expressed as

$$E = \frac{1}{2}\dot{\mathbf{q}}^T I \dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^T K \mathbf{q} \quad (3)$$

where $\tilde{K} = \text{diag}\{\omega_i^2\}$. Differentiating equation (3) with respect to time and substituting in equation (2), the energy dissipation rate can be expressed as

$$\frac{dE}{dt} = -\dot{q}^T \tilde{D} \dot{q} \quad (4)$$

where $\tilde{D} = \text{diag}\{2\zeta_i \omega_i\}$. Integrating equation (4), the dissipation energy can be evaluated as

$$E_d = \int \dot{q}^T \tilde{D} \dot{q} dt \quad (5)$$

$$= \int z^T \begin{bmatrix} 0 & 0 \\ 0 & \text{diag}\{2\zeta_i \omega_i\} \end{bmatrix} z dt$$

Note, since the dissipation energy is a negative quantity, the negative sign was omitted in equation (5) for convenience. Using the solution from the state space representation of the closed-loop system, $z = e^{At} z_0$ results in the integral

$$E_d = z_0^T \int e^{A^T t} \begin{bmatrix} 0 & 0 \\ 0 & \text{diag}\{2\zeta_i \omega_i\} \end{bmatrix} e^{At} dt z_0 \quad (6)$$

For optimal placement, equation (6) can be partitioned into a contribution from inherent structural damping and the addition of passive members (see [CHEN89]), i.e.,

$$E_d = E_i + E_{pj}$$

The energy dissipation rate is maximized over a finite period of time, about five times the maximum period in the system.

The computational complexity of a combinatorial problem can be expressed in the following form

$$n = \frac{I!}{I^*! (I-I^*)!}$$

where I = number of structural members and I^* = number of passive members. Clearly, the number of possible combinations becomes rather large even for structures of small magnitude. Therefore, instead of performing an exhaustive combinatorial optimization, a heuristic based approach was used that renders near optimal solutions in a less extensive search. When finding an approximate solution, the most common technique used is an iterative improvement method where only improvements in the performance criterion are accepted. This type of optimization has a tendency of getting trapped at local optima, and without any means of climbing away from these local values, a more global value cannot be found. Simulated annealing is a modified iterative improvement technique that provides a method which avoids becoming trapped at local optima.

The simulated annealing technique (SA) was developed by Kirkpatrick et al [KIRK83], when he observed a correlation between combinatorial optimization and the annealing process of solids. Using the Metropolis algorithm that numerically simulates the behavior of atoms in a solid at a finite temperature, Kirkpatrick developed an equivalent method for use in a combinatorial problem. SA allows the acceptance of non-improving solutions by using a probability function $P = e^{\Delta E/k\theta}$, where ΔE is the change in energy, k is a normalized value and θ is a free parameter that governs the acceptance of non-improving solutions. Therefore, SA allows the program to achieve a more global solution than is possible by an iterative improvement method. A flow chart of the simulated annealing algorithm is shown in Figure 1. The schedule that is chosen to decrement θ has a direct affect on the convergence of the solution to a global optimum. Subsequent work will consider the use of a nonlinear schedule that is tailored to the study of the twenty bay truss structure. The generation of possible configurations of the passive members was accomplished by randomly moving the passive members one at a time, then the difference between the old configuration's cost function and the new configuration's cost function was calculated. Since the performance criterion was to be maximized, only a positive

change in the cost function was accepted. When a non-improvement occurs, the probability function was calculated and compared to a random number. If the random number was less than the probability function, the configuration was accepted and the program updates the configuration. As θ decreases, the likelihood of accepting non-improving solutions becomes small, and near the completion of the optimization the program becomes a purely iterative improvement based technique.

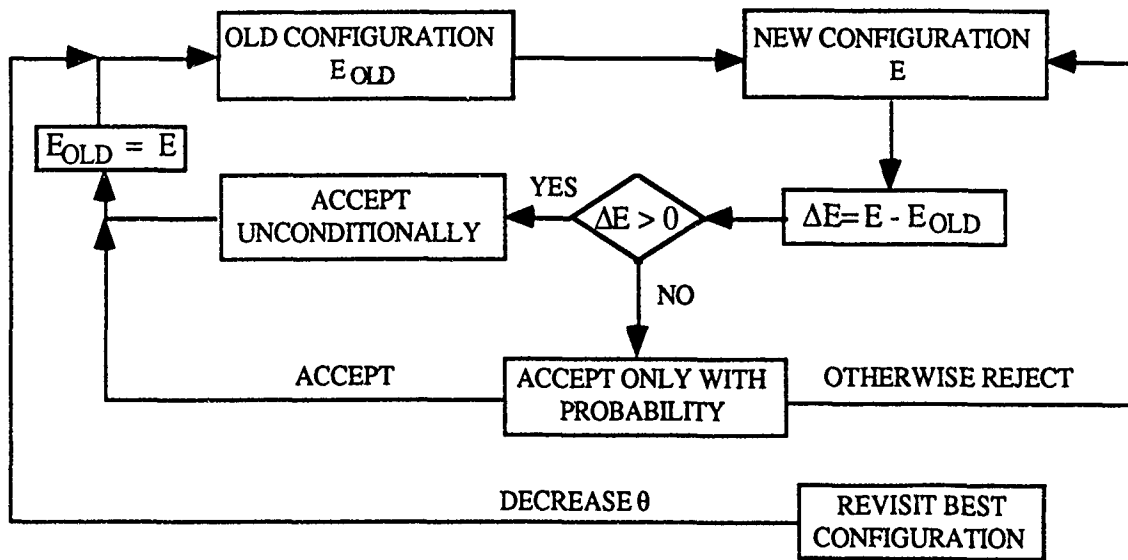


FIGURE 1. Algorithm of Simulated Annealing*
(* taken from [CHEN89])

Results

This study was conducted on an 80 member twenty bay truss structure. Only sixty locations in the truss were considered for passive member placement. The third mode of vibration of the truss was the only mode examined because the results could easily be compared to the method of modal strain energy (MSE). Basically, MSE determines the

structural loss factor associated with the i th mode and yields members that are the optimal locations for placement of a finite number of passive dampers. For the third mode, MSE found the global solution for four passive members to be located at [1 3 53 55]. Figure 2 shows these locations in the schematic of the truss.

Simulated annealing uses parameters that are structure dependent. The initial value of ΔE was found by averaging the cost function over twenty random combinations. This value of ΔE was then used to find the initial value for the annealing parameter θ , where $\theta = \frac{\Delta E}{\ln(P)}$. A probability factor of .8 was used which gives an initial acceptance rate of 80% for non-improving solutions. The annealing parameter was then reduced by a constant factor $\alpha = .9$. Note that this gave a linear optimization schedule for a nonlinear problem. Continuing work will use a nonlinear schedule to enhance the convergence of the solution.

Figure 3 shows the iteration history of a typical solution for four passive members when examining the third mode. This graph demonstrates how the program converges to a near optimal solution. It can be seen that towards the end of the trial, few non-improving solutions were accepted. The dashed line represents the optimal E_d value retained while searching random combinations. A comparison was made between simulated annealing and iterative improvement to show the advantage of this technique. Figure 4 shows that the simulated annealing algorithm obtained a higher energy dissipation rate than the iterative improvement method. It clearly shows that simulated annealing has the ability to climb away from a local optimum value. Both methods were started with the same initial combination of passive member placement.

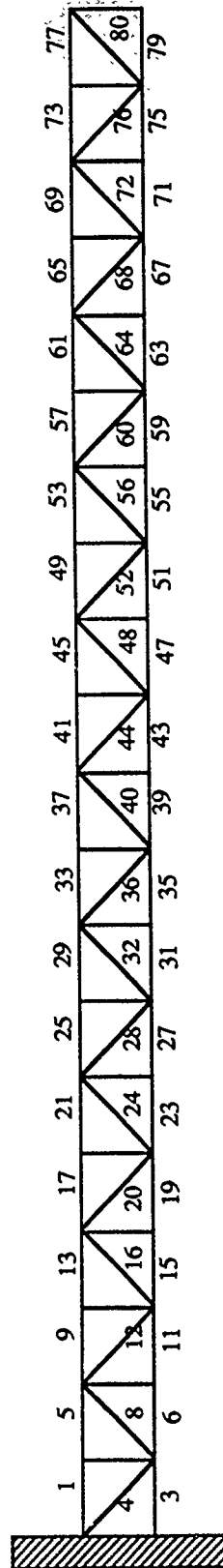


Figure 2. 20 bay planar truss with location numbers

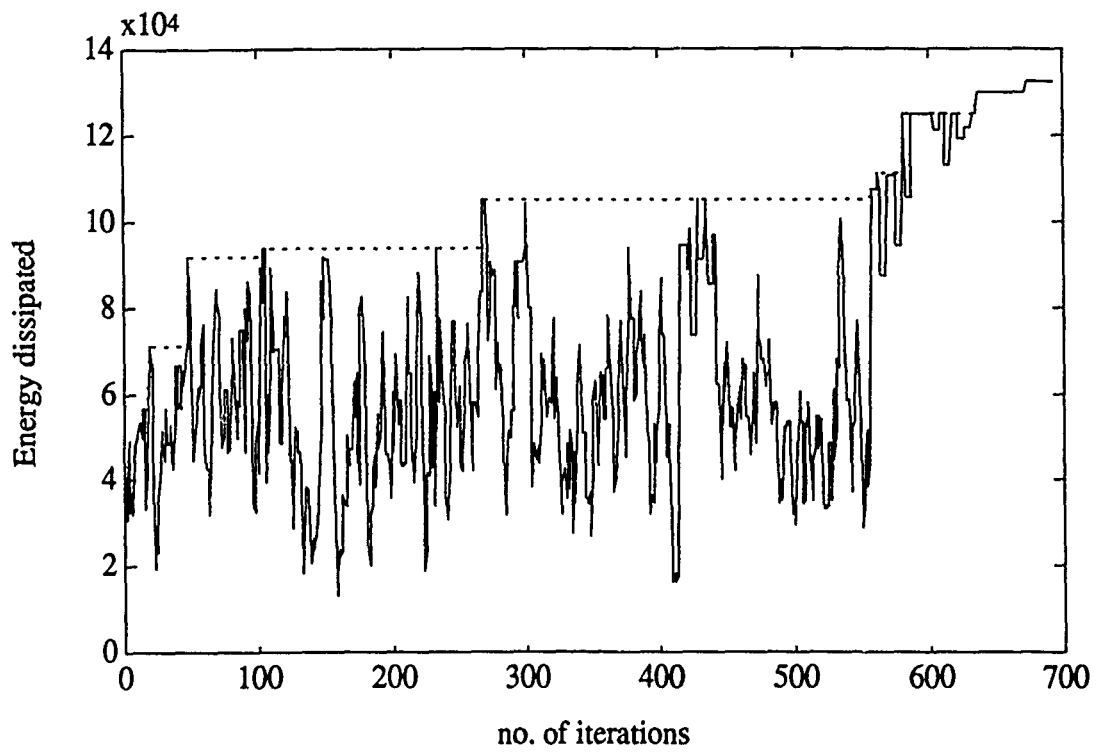


Figure 3. Iteration history of the simulated annealing algorithm

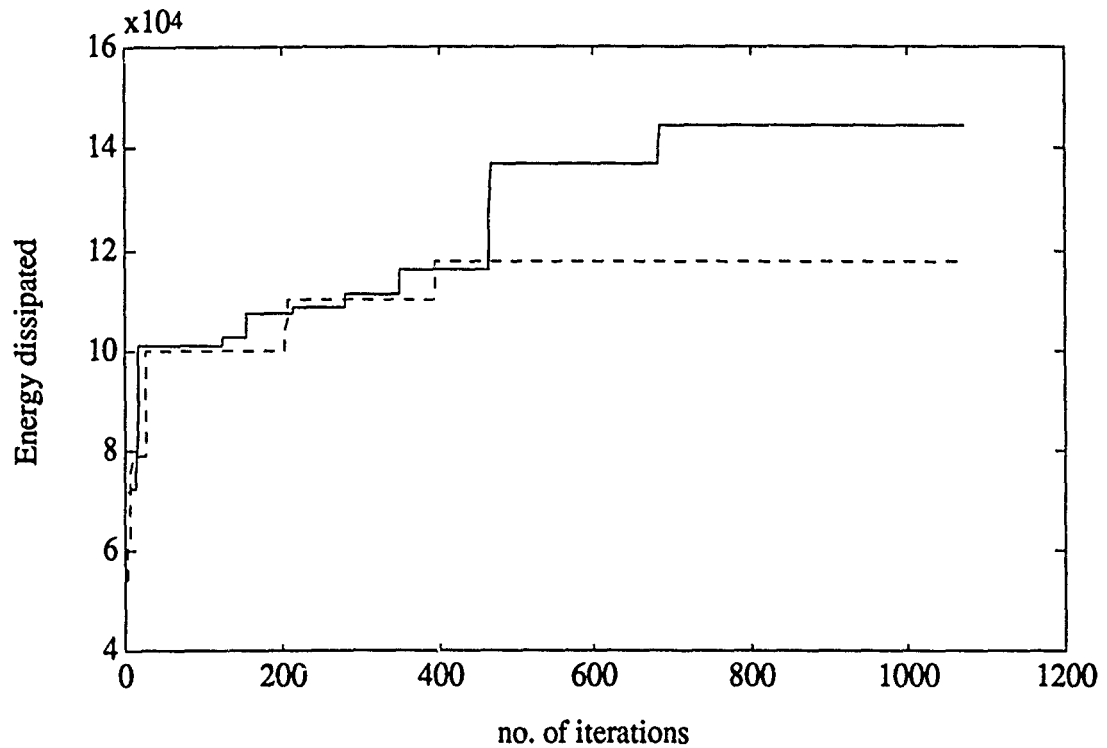


Figure 4. Comparison of Simulated Annealing and Iterative improvement

Various trials were conducted to compare the solutions from different starting combinations. Referring to Figure 5, and noting that the number of iterations depends on the initial value of θ , it can be seen that SA continually yields near optimal solutions. The trial with the highest cost function value gave a combination of [1 3 55 57] as the optimal solution. The dashed line on the graph represents the energy dissipated due to the inherent structural damping. Thus, this graph also shows how the addition of four passive members greatly enhances the damping of vibrations in this structure

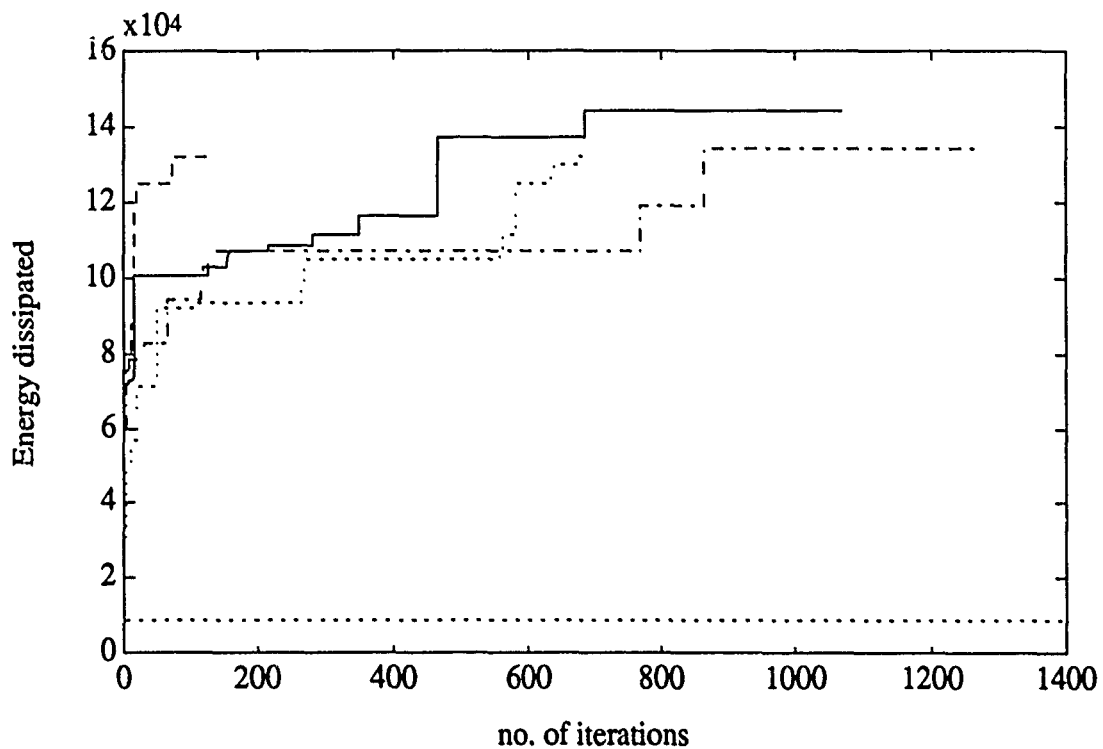


Figure 5. Comparison of near optimal solutions

Conclusions

Simulated annealing was shown to provide a means of obtaining near optimal solutions for passive member placement. This study was conducted on the third mode of vibration of the truss structure. The results of this study show that SA yields near optimal solutions while evaluating less than .3% of all possible combinations. Continuing work is being done to enhance the convergence of the algorithm and then examine the convergence of the SA technique for a cluster of modes.

Acknowledgement

This study was conducted at the Frank J. Seiler Laboratory located at the US Air Force Academy. This work was sponsored by the Air Force Office of Scientific Research.

References

- [CHEN89] Chen, G.-S., Bruno, R. and Salama M., "Selection of Active Member Locations in Adaptive Structures", *Proceedings of the 30th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Mobil AL, April 1989, AIAA Paper 89-1287.
- [KIRK83] Kirkpatrick, S., Gelatt, C. and Vecchi, M., "Optimization by Simulated Annealing", Science, Vol. 220, No. 4598, May 1983, pp. 671-680.
- [SALM88] Salama, M., Bruno, R., Chen, G.-S. and Garba, J., "Optimal Placement of Excitations and Sensors by Simulated Annealing," Second NASA/Air Force Symposium on Recent Experiences in Multidisciplinary Analysis and Optimization, Hampton VA, Sept. 1988, NASA-CP-3031-PT3.

SECOND HARMONIC GENERATION IN POLED AMORPHOUS QUARTZ

Mr. Karl E. Kauffmann

Abstract

In this paper the observance of second harmonic generation from poled amorphous quartz is reported. Samples are heated and then poled under high field strengths. These field strengths are maintained upon subsequent cooling, and the second harmonic signal of incident 1064 nm light is observed using the Maker fringe technique. The dependence of field strength on the magnitude of the second harmonic is explored. The temperature dependence on the poling process as well as that upon subsequent heating of the poled samples is studied. The observed second harmonic is shown to be relatively stable for several months at room temperature. ESR studies have been performed to determine the role of paramagnetic centers in the generation of the second harmonic signal. The nature of the dipoles aligned in the poling process has not been determined.

Introduction

Ever since the demonstration of second harmonic generation (SHG) by Franken in 1960, nonlinear optics has established itself as a field of discipline for many scientists. Today, it is one of the youngest and most active areas of research.

The reason for such high interest in this field lies in the potential application of nonlinear optics (NLO) for use in communications as well as computer hardware design. In the past twenty years or so, the field of electronics, especially with the advent of computers, has exploded into one of the largest areas of research and development. However, due to the actual physical processes involved in the passage of electrons through an electronic device, there are inherent limitations on the space, efficiency, and speed with which the device may operate. To overcome these problems, researchers are turning to

the idea of using light to perform those same functions. The idea of an all-optical device, like a microprocessor chip, functioning at the speed of light with minimal loss is quite an appealing direction for modern day technology to head.

Presently, research has been directed at finding materials which yield relatively high optical nonlinearities while at the same time exhibiting fast response to an applied field. These materials also must show an ease in fabrication, high optical quality, and high resistance to damage.

Though this direction of research is a noble one, much of the research exploring the physical mechanisms, leading to NLO behavior have not been explored. As such, there is still much that is unknown about the nonlinear optical response mechanisms of various systems. For example, the polarization in the third order response of the uniformly aligned poly (p-phenylene benzobisthiazole) system does not correspond with the optical axis of the sample.¹ Why? Essentially, this suggests that the large third order response of the PBT system occurs as a result of π -delocalization over a correlation length much smaller than would be expected.

Another quite interesting example is the fiber optic "puzzle." If an incident laser beam of 1.064μ is coupled into a Ge and/or P doped optical fiber, after a length of time the second harmonic will be generated by the medium even though symmetry conditions of the fiber would suggest that this process is forbidden. Many theories attempting to account for the phenomenon have been proposed,² but none seem to accurately account for all observations of the behavior of these systems.

Originally, the project discussed in these pages started out as a simple characterization of a nominal second order nonlinear response observed in poled amorphous quartz. We had hoped to quickly generate a paper based on our observations. However, the project became much more complex than was first thought yielding results that we could not seem to piece together to say that we actually knew what was going on in the system. The goal of this project, then, was to not only characterize the NLO properties of the poled amorphous quartz system but to also explain the actual physical mechanisms giving rise to the observed nonlinearity. We believe that the problem may relate to that of the fiber optic system. Thus, in-depth studies of this phenomenon should yield considerable insight to the processes involved in the SHG observed in doped fiber optic cable.

Theory

Normally when monochromatic light passes through some dielectric that is optically clear, i.e. nonabsorbing at the given wavelength, we expect that we will see the same light coming out as that going in. We learn this from everyday experience. However, even though this is generally true, it is not always the case. If the light incident on an optically clear medium is sufficiently intense, it can induce nonlinearities in the material that will in turn alter the exiting light. The field of nonlinear optics is devoted to studying such effects and their mechanisms.

When light enters into an optically clear dielectric, it induces polarizations in the medium that in turn propagate the light through the material. A nonlinear system can be characterized by an expansion of the polarization term, \mathbf{P} , for the medium. This expansion yields:

$$\mathbf{P}_i = \mathbf{P}_i^0 + \chi_{ij}^{(1)} E_j^\omega + \chi_{ijl}^{(2)} E_j^\omega E_l^{\omega'} + \chi_{ijlm}^{(3)} E_j^\omega E_l^{\omega'} E_m^{\omega''} + \dots$$

where \mathbf{P}^0 is the permanent polarization in the medium and $\chi^{(n)}$ is the n^{th} order susceptibility.

For light having low \mathbf{E} field intensities, $\chi_{ij} E_j^\omega \gg \chi_{ijk} E_j^\omega E_k^{\omega'}$. As a result, the higher order terms can be ignored, and a linear relationship is observed between \mathbf{E} and the polarization in the medium. If $\mathbf{P}^0 = 0$, then

$$\mathbf{P} = \epsilon_0 \chi^{(1)} \mathbf{E}.$$

However, when the \mathbf{E} field intensities become sufficiently large (on the order of 10^7 V/cm), the higher order terms can no longer be ignored, and therefore, the relationship between \mathbf{P} and \mathbf{E} becomes nonlinear in behavior. Since there are numerous effects that can occur in considering higher order χ 's, discussion will be limited to the evaluation of the $\chi^{(2)}$ term (Even for just the $\chi^{(2)}$ term alone there are several effects due to the mixing of two fields that can be observed). It is important to note that all $\chi^{(2)}$ effects require the material to have a noncentrosymmetric structure. If the material exhibits inversion symmetry, then $\chi^{(2)} = 0$.

In studying $\chi^{(2)}$ effects the flow of energy through the material is of primary concern. This is necessary because the direction of much of the research in this field is to maximize the efficiency of the conversion of the two input fields and

thereby, maximize the nonlinear effect so that it can be used in a practical application.

From the conservation of energy,

$$\omega_3 = \omega_1 + \omega_2$$

where ω_3 is the frequency of the nonlinearly generated wave, and ω_1 and ω_2 are the frequencies of the input fields.

The phase difference of the waves travelling through the medium is given by

$$\Delta k = k_3 - k_1 - k_2$$

where k_i is the propagation constant for the i^{th} wave in the medium.

Avoiding much of the mathematical rigor, the derived equation for the flow of energy in the material in the low conversion limit is:³

$$S_3 = (\omega_3^2 / 2n_1 n_2 n_3) (\mu/\epsilon_0)^{3/2} d^2 L^2 S_1 S_2 (\text{sinc}^2(\Delta k L / 2))$$

where S_i is the magnitude of the Poynting vector (energy flux), n_i is the refractive index for the i^{th} wave in the medium, L is the distance travelled into the medium, and d is the nonlinear optical coefficient that is experimentally determined ($d \propto \chi^{(2)}$).

One aspect about the above equation which should be pointed out is that in the case of perfect *phase matching*, i.e. $\Delta k = 0$, S_3 should grow as L^2 since the $\text{sinc}^2(\Delta k L / 2)$ term will be equal to one. Usually, though, perfect phase matching is not achieved and the intensity of the generated wave will vary sinusoidally through the medium. Therefore, a distance l_c into the medium is defined for which S_3 is a maximum. This can be found by taking the first derivative of the energy flux equation with respect to L and then setting this derivative equal to zero. The coherence length, l_c , is then defined as⁴

$$l_c = |\pi / \Delta k|.$$

In performing nonlinear optical measurements, one must pay special attention to the magnitude of the path travelled by the generated wave in the medium for reasons of phase mismatch.

For second harmonic generation, $\omega_1 = \omega_2$ and the generated frequency, ω_3 , is twice that of the incident wave. The above equations then become:

$$\omega_3 = 2\omega_1,$$

$$\Delta k = k_3 - 2k_1, \text{ and}$$

$$S_3 = (\omega_3^2 / 2n_1^2 n_3) (\mu/\epsilon_0)^{3/2} d^2 L^2 S_1^2 (\text{sinc}^2(\Delta k L / 2))$$

where $d = \chi^{(2)} / 2$.

For more information on nonlinear effects in optics, one can consult any of several texts on the subject.^{3,4,5,6,7}

Experimental

Theory

The experiment used to measure the second harmonic intensity was the same as that used by Maker *et al.*⁸ and is coined as the "Maker fringe technique." Since the theory for this experiment is rather involved, only the basic concepts will be presented here.

In a nutshell a laser is focused on a plane parallel sample that is flat to within a wavelength of the incident beam. The electric field, E_ω , of the focused beam induces nonlinear polarizations in the sample. The waves travelling through the medium fall into one of two categories. At the surface a "free wave" is formed that passes through the sample with an effective index of refraction $n_{2\omega}$. Within the sample a "bound wave" is formed that travels with a refractive index n_ω . Since these free and bound waves move through the sample at different velocities (see figure 1), interference occurs at the exit of the sample giving rise to fringes as the sample is rotated about an axis. As a result, the harmonic power, $P_{2\omega}$, can be analyzed as a function of the incident angle, θ . A Maker fringe pattern for quartz is shown in figure 2.

From calculations using Maxwell's equations, the final transmitted second harmonic power can be written:⁹

$$P_{2\omega} = d_{2\omega}^2 (512 \pi^3 / w) t_\omega'^4 T_{2\omega}''^4 \rho(\theta) \{ \rho(\theta) \}^2 \beta(\theta) P_\omega^2 \{ n_\omega^2 - n_{2\omega}^2 \}^{-2} \sin^2 \Psi,$$

where $d_{2\omega}$ is the NLO coefficient, w is the Gaussian beam width incident on the sample, t_ω' and $T_{2\omega}''$ are the transmission factors of the front and back faces of the sample, respectively, $\rho(\theta)$ is a correction for multiple reflections at the air-medium/medium-air interfaces, $\beta(\theta)$ is a correction for the beam width, $\rho(\theta)$ is a projection factor that varies depending on the symmetry of the medium being analyzed and its orientation with respect to the incoming fundamental, P_ω is the power of the fundamental, and Ψ is defined by:

$$\Psi = (\pi D / 2) (4 / \lambda) \{ n_\omega \cos \theta_\omega - n_{2\omega} \cos \theta_{2\omega} \}$$

For the purposes of discussion, the details of the above equation need not be considered in full. Therefore, the above equation can be written in a simplified form

$$P_{2\omega} = C \{d_{2\omega}^2\} \{ENV(\theta)\} \sin^2\Psi,$$

where C is a constant,

$$C = (512 \pi^3/w) P_{\omega}^2,$$

and ENV(θ) is the envelope function,

$$ENV(\theta) = t_{\omega}^4 T_{2\omega}^4 p(\theta) (p(\theta))^2 \beta(\theta) \{n_{\omega}^2 - n_{2\omega}^2\}^{-2}.$$

Looking back at figure 2, the C and $d_{2\omega}$ terms determine the amplitude of the oscillations in the Maker fringe pattern. The envelope function in turn is responsible for the beat-like appearance of the spectrum and characterizes the magnitude of the fringe maxima. The $\sin^2\Psi$ term determines the individual oscillations of graph.

Experimental Setup

The design of the apparatus used to perform the Maker fringe experiment is diagrammed in figure 3.

The choice of this system is, by far, not unique. Essentially, the reason for this design as opposed to any others was for convenience of space. It was originally believed that a third harmonic generation setup could be run off the same laser used to generate the second harmonic signal. However, it was later discovered that the pulse rate of the laser was too great to be put in sync with a Raman shifter. Therefore, the designing of a third harmonic system during the course of my ten week project was eventually abandoned since a pulse laser having the proper rate was not available for use. The original setup for the Maker fringe experiment was kept for mere convenience.

During the first two weeks of study, attempts were made to optimize this system. Several characteristics of the system were found to exist that affected both the power and the quality of the beam. Over the course of the first half of this project, most of these quirks were eliminated by proper location of the front reflecting prism, positioning of available lenses, and realignment of the photomultiplier tube. The system, however, still needs to be readjusted in terms of its choice of lenses. With the present lens system the focal point on the sample is too "tight." As a result, it heats the sample while the experiment is being run causing minute changes in the sample's refractive index. The effect

on the Maker fringe pattern is to cause the oscillations to shift toward the zero degree angle since small changes in the refractive index are reflected in the coherence length of the sample. Our solution to this problem was to adjust the laser power to a low enough level that sufficient heating could not be achieved (I might add that by optimizing the system the amount of power required to achieve a comparable signal to those previously measured was cut by a factor of ten!).

Sample Preparation

The samples studied in this project were CO quartz cover slips made by Esco Products Inc. (Part number R525000). The standard used was a y-cut quartz disc about 0.240 cm thick.

Samples were first measured for thickness using an electronic thickness gauge. Samples were on the order of 150 to 200 μ . However, the faces of the samples were not what one could call very parallel. Thicknesses over the sample surface could sometimes vary by as much as 10 μ . This probably accounts for some of the unique symmetries exhibited by the SHG fringe patterns for the samples that had been poled and analyzed (The Maker fringe technique assumes that the faces of a sample are parallel to within a wavelength of the fundamental). Since the variations in the sample thickness could vary by relatively large amounts, the thickness was taken as an average over many points in a given sample.

Once the thickness of a sample was known, it could then be heated and poled. Typical applied fields ranged from 3000-7000 V. The lower limit was obviously set by the sample and our ability to view the SHG phenomenon. The upper value, however, resulted from an inherent flaw in the design of the poling cell. At such large fields the current could arc around the sample which could cause major damage to the power supply. Also above this voltage, it became very easy to cause physical macroscopic damage to the sample being poled.

The sample was cooled over a slow period of time while maintaining the field across the sample. Once the sample reached room temperature, the electric field was turned off, and the sample was removed for analysis.

A usual heating cycle consisted of the following steps: 1) Heat the sample to 250°C; 2) Apply the D.C. field; 3) Allow the sample to cool to 200°C over a period of two hours; 4) Allow the sample to cool to 150°C for about an hour; 5)

Allow the sample then to further cool to 100°C for about an hour; and 6) Cool to room temperature. Samples quenched did not exhibit significant amounts of SHG. This would explain the need for a slow annealing process as described here. Variations on this procedure were tried with varying amounts of success; however, this particular procedure appeared to provide the best results.

Some samples were also γ -irradiated with a ^{60}Co source in order to generate a large number of defects within the structure of the amorphous quartz. Samples were then poled according to standard procedures as outlined above.

Other Studies

Several other studies were done on the samples to determine structure and the nature of this "pseudo-transition" which gave rise to the second harmonic. Several instruments were called upon to perform these tasks. These analyses made use of scanning electron microscopy (SEM), energy dispersive spectroscopy (EDS), electron spin resonance (ESR), and differential scanning calorimetry (DSC).

Results

The primary goal of this ten week project was to gather as much possible data as could be collected on this observed SHG effect. Explanation of the center responsible for the SHG became necessary only upon the realization that paramagnetic defects did not appear to be the cause for the SHG. No $\langle d \rangle$ values have been calculated yet. We are awaiting the completion of a computer program to calculate these from the numerous spectra we have now available to us (I care not to report premature results until a methodology is established by which we can analyze the available data).

The Maker fringe pattern for these samples were similar to those obtained for poled polymer samples.¹⁰ A typical spectrum is shown in figure 4.

From studies done on the relationship between the poling field and the relative magnitude of the SHG signal, it can be seen that the SHG effect increases in magnitude with increased field strength (see figure 5). The exact relationship between the poling field and the observed SHG has still not been determined. It may be that the curve should be linear having a minimum cutoff

frequency, and our errors due to laser realignment has caused a substantial variance in the data. It has been suggested by Dr. Marek Grabowski that the data may suggest an exponential-type fit which would correspond to the fields ability to break the symmetry of the film. In playing with various curve fitting programs the data appears to correlate best to a third order polynomial which may suggest the possibility that our field strengths are sufficiently high ($O(10^5)$) to induce contributions from nonlinear effects.

Several experiments to observe relaxational effects of the dipoles in the samples were performed. The induced order of the system has been shown to be quite stable for several months (see figure 6). In terms of these studies, the "hump" that appears in this graph suggests an interesting effect going on within these films a few days after poling. In some of the films, the relative SHG was observed to increase within a time period of about a week after poling. This may be indicative of relaxations occurring in the sample that serve to actually induce more ordering in the systems and thereby increase the observed SHG.

Upon subsequent heating the induced effect has been shown to be stable up to 130-150°C. The SHG then drops off at higher temperatures (see figure 7). The poled films are also observed to pass a current upon heating. Current vs. temperature studies have been performed which show that as the poled samples are heated a current on the order of 10^{-9} amps is produced by the samples in a range of 150-250°C reaching an apparent maximum somewhere above 200°C (which may explain why samples poled at temperatures lower than 200°C exhibited lower SHG behavior than samples poled by the normal procedure under comparable field strengths). This implies that dipoles are being aligned or created as a result of our preparation procedure. Thus, the observed loss in the SHG occurs as a result of the loss of a noncentrosymmetric center and/or dipole annihilation.

Variations on the preparation procedure were also investigated. Sample preparation was performed in an inert atmosphere. Since the SHG was still observed, it allowed us to rule out the possibility that contaminants from the air were being introduced into the films. Sample preparation was also done whereby a film was effectively shielded from the electrodes using PES. In this case the observance of the SHG signal allowed us to rule out injection of charged ions by the electrodes used to pole the samples. From these two

experiments we have concluded that the observed SHG is due to some species unique to the a-SiO₂ films.

ESR studies were performed on poled and unpoled samples to determine the role of paramagnetic defect centers in the SHG process. Samples irradiated with γ -radiation from a ⁶⁰Co source were poled and then analyzed using both the Maker fringe experiment and ESR. ESR confirmed that, in fact, the irradiated films did exhibit an increased concentration of defects. However, the lack of enhancement to the observed SHG has led us to believe that no correlation exists between the paramagnetic defects common in quartz and the dipoles which are being created by the poling procedure. This was further confirmed by a sample that was annealed at 700°C prior to poling. This sample exhibited SHG behavior but showed no detectable paramagnetic centers from ESR.

Poled samples were also analyzed using DSC. The results of this study showed no type of phase transition going on in the temperature window (<250°C), which was expected. SEM analyses were performed to see if any macroscopic structural defects could be observed in the poled samples, and as expected no imperfections in the surface of the samples could be observed. Subsequent electron emission studies showed no detectable impurities in the poled films. However, they did show that there was an unusually high percentage of oxygen at the film interface. Whether or not this is indicative of anything will rely on further experimentation.

Attempts at inducing a $\chi^{(2)}$ grating by a seeding process using phasematched 532 nm light with the fundamental were made with no success. We should not, however, rule out the possibility that the fiber optic problem and what is being observed here are related.

Discussion

Initially, when this project was started, the observed SHG was believed to be due to an aligning of paramagnetic defect centers in the samples. This belief, quite logically, stemmed from the assessment that in doped fibers the structural unit responsible for the observance of SHG is a GeE' center.¹¹ However, the data collected from annealed and irradiated samples have essentially "blown

this theory out of the water." Therefore, we have yet to come up with a sufficient physical model to account for the phenomenon we are observing.

We do, though, propose two different possibilities for the observance of the SHG in these films. Eventually, each one of these routes will be explored and will either be incorporated or discarded depending on the results of further experimentation.

1) *Ion Mobility/ Dipole Alignment*

a) Impurities

One of the ideas is that impurities in the quartz films may be the culprits behind the observance of the SHG. This does have some support since it is believed that impurities in the fiber optic systems give rise to the SH signal. There is, however, no germanium reported to be in these films. Since this is the case, then the effect must be due to some type of center other than the GeE' defect.

One proposed theory is that an $[\text{AlO}_4]^-$ defect could give rise to an effective dipole by having its charge compensator(s) moved a distance away from the $[\text{AlO}_4]^-$ center. It is known that above 200°C sodium and potassium cations become mobile along the z-axis in a quartz crystal.¹² This is simply due to the influence of thermal energy on the crystalline structure of quartz. At elevated temperatures, the volume surrounding the z-axis of the quartz crystal expands allowing for the mobility of small ions trapped within the crystal. In the case of our preparation process, the z-axes of the microcrystallites comprising the samples could be expanded allowing for the mobility of Na^+ and/or K^+ along the axes. Therefore, in the presence of an external field, these cations could be preferentially directed to break the symmetry of the system and thus give rise to an SHG signal.

The argument against such a theory is that the concentration of aluminum, or for that matter any impurity, in the samples may be too low (see Table 1) to yield an effective concentration of dipoles in the system to give rise to an observable SHG.

b) Other Si Defects

Another question may be raised as to what kind of defects could give rise to effective dipoles in the films. Most of the literature sources on defects in quartz

have analyzed only paramagnetic centers. None have analyzed the possibility that other types of defects may be present in quartz that are not capable of being probed using ESR techniques. In two of the sources read, it is mentioned that there are precursors to the E' and E'' states that are formed upon the heating of quartz between 200 and 300°C.^{13,14} The structure of these precursors, however, have not been identified since they are not paramagnetic.

2) Charge Trapping

This is probably the most vague of the proposed ideas. Its basis lies in the fiber optic problem.

There are essentially two competing schools of thought on the nature of the physical mechanism giving rise to the SHG in doped fibers. Both involve germanium but are based on different premises. In the one case it is believed that the SHG is due to an actual sinusoidal alignment of Ge defects in the fiber.² This theory, however, lacks coherency simply due to one major argument. It is not believed that the energetics involved in breaking the symmetry of the fiber are sufficient to cause actual physical alignments of dipoles in the system. The other theory, which is gaining ground in the community that is studying this problem, espouses that electrons as a result of a four photon absorption process are ejected from Ge sites in preferential directions.^{15,16} These electrons then become trapped at other sites in the fiber giving rise to a photorefractive grating that permits the growth of the second harmonic. The only thing not mentioned in this theory is the nature of the trap sites, which is obviously unknown.

In our case it is possible that due to the D.C. poling field that is being applied to the quartz films in addition to the thermal energy, a current is being passed through the samples which allows electrons to move about in the films. Upon cooling, these charges become trapped at preferential sites which in turn break the symmetry of the α -SiO₂ films. The nature of these traps are about as well defined as those proposed in the fiber optic problem.

One or all of these ideas may apply to the phenomenon reported here. It could be that some impurity or Si defect (other than a paramagnetic center) gives rise to the SHG by charge ejection and/or trapping. It could be also possible that we are "all wet" in the belief that any of these mechanisms could

give rise to the observed SHG. Just the same, we should not rule out any of them until direct proof of what is happening in these systems due to our preparation procedure has been established.

Conclusions

In summary, we have observed SHG in thin films of quartz that have been prepared by simultaneous heating and poling. The magnitude of the SHG has been shown to depend on both the strength of the poling field and the poling temperature(s) used in the preparation of the thin films.

The stability of the poled films with respect to time and/or subsequent heating has been investigated. Also, current vs. temperature studies have been performed on the samples during the poling process and during subsequent heatings. These studies have shown that effective dipoles are being aligned and/or created in these films giving rise to the observed SHG.

Variations in the preparation process such as heating and cooling in a nitrogen atmosphere and sample shielding have been performed to eliminate the possibility that the SHG is due to contamination from the poling process.

ESR studies have been done on poled and unpoled samples to determine the role of paramagnetic centers in the observance of the SHG phenomenon. The effects of γ -irradiation on the enhancement of the observed SHG have been investigated.

DSC analysis of poled samples have been performed showing no observable phase transition as a result of relaxation of the "dipoles" aligned in the sample. SEM scans of the samples have also been done to observe any macroscopic indications of induced order. Electron emission of the samples have indicated no abnormally high impurity in the samples.

From the results of this ten week project, it appears that we have raised more questions than anticipated. What appeared to start out as a simple optical characterization of an SHG phenomena in a poled quartz system has ended up as one big puzzle. We have not determined what the nature of the center is in the system that gives rise to the SHG. From our studies, we have basically done more to rule out possibilities than to actually "rule them in." Anything that has been discussed in this paper relating to what the center may be has been

simply based on conjecture. More experimentation of the phenomenon needs to be done in order to determine the true nature of the dipoles in the system.

Acknowledgments

I would like to thank Research Development Laboratories (RDL) and the United States Air Force for sponsoring my research work over this ten week period. I would also like to extend sincere gratitude to Dr. John Kester whose laboratory, time, and patience were most appreciated and a special thanks to all the people of the laser physics division of the Frank J. Seiler Research Laboratory who made my summer not only insightful but also fun.

References

- 1) H. Mattoussi and G.C. Berry; *The Third Order Susceptibility of Nematic Solutions of PBT*; unpublished; April, 1991.
- 2) R.H. Stolen; *Second Harmonic Generation in Optical Fibers*; preprint to a published article; July 1989.
- 3) R. Guenther; *Modern Optics*; John Wiley & Sons, New York; 1990.
- 4) P.W. Milonni and J.H. Eberly; *Lasers*; John Wiley & Sons, New York; 1988.
- 5) Y.R. Shen; *The Principles of Nonlinear Optics*; John Wiley & Sons, New York; 1984.
- 6) A. Yariv; *Optical Electronics: Third Edition*; Holt, Rinehart, & Winston, New York; 1985.
- 7) P.N. Butcher and D. Cotter; *The Elements of Nonlinear Optics*; Cambridge University Press, Cambridge; 1990.
- 8) P.D. Maker *et al.*; **Phys. Rev. Lett.**; 8; 21 (1962).
- 9) J. Jerphagnon and S.K. Kurtz; **J. Appl. Phys.**; 41; 1667 (1970).
- 10) K.D. Singer *et al.*; **Appl. Phys. Lett.**; 49; 248 (1986).
- 11) T.E. Tsai *et al.*; *Topical Meeting on Nonlinear Guided-Wave Phenomena: Physics and Applications, Technical Digital Series*; Vol. 2 (p. 250); Optical Society of America, Washington D.C.; 1989.

- 12) M.E. Markes and L.E. Halliburton; **J. Appl. Phys.**; 50; 8172 (1979).
- 13) M.G. Jani *et al.*; **Phys. Rev. B**; 27; 2285 (1983).
- 14) R.B. Bossoli *et al.*; **Solid State Comm.**; 44; 213 (1982).
- 15) N.B. Baranova and B. Ya. Zel'dovich; **J. Opt. Soc. Am. B**; 8; 27 (1991).
- 16) D.Z. Anderson *et al.*; *A Model for Second Harmonic Generation in Glass Optical Fibers Based on Asymmetric Photoelectron Emission from Defect Sites*; preprint to a published article; December, 1990.

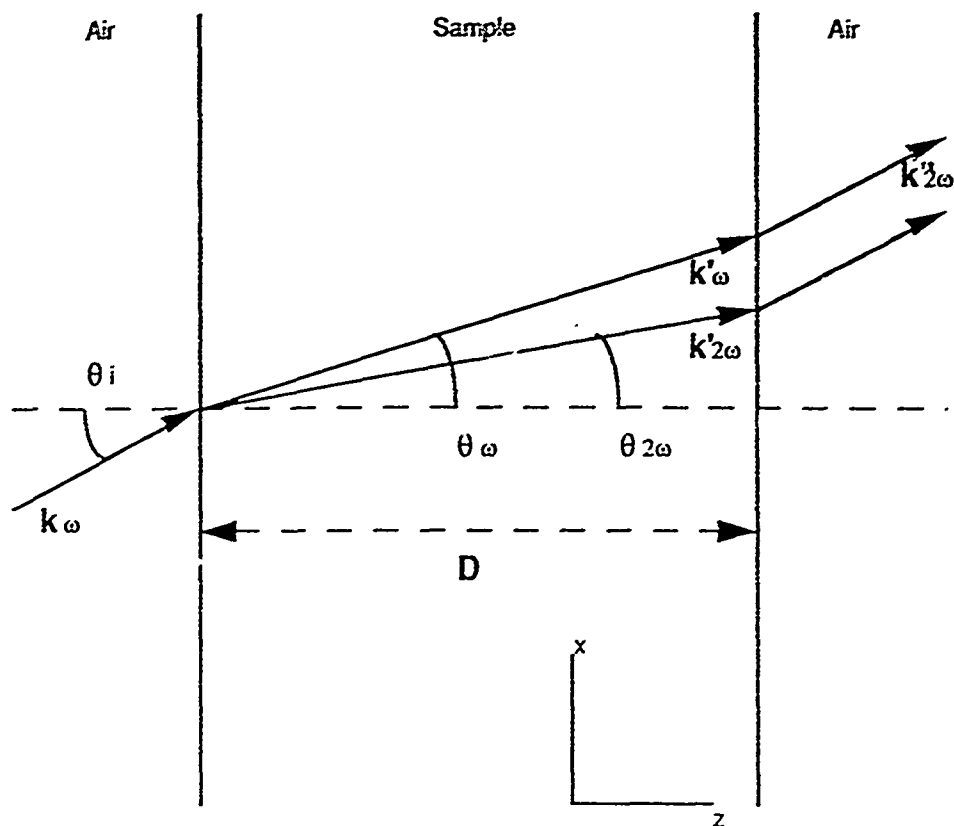


Figure 1. Diagram of wave propagation in a dielectric showing both the bound and free waves travelling through the medium.

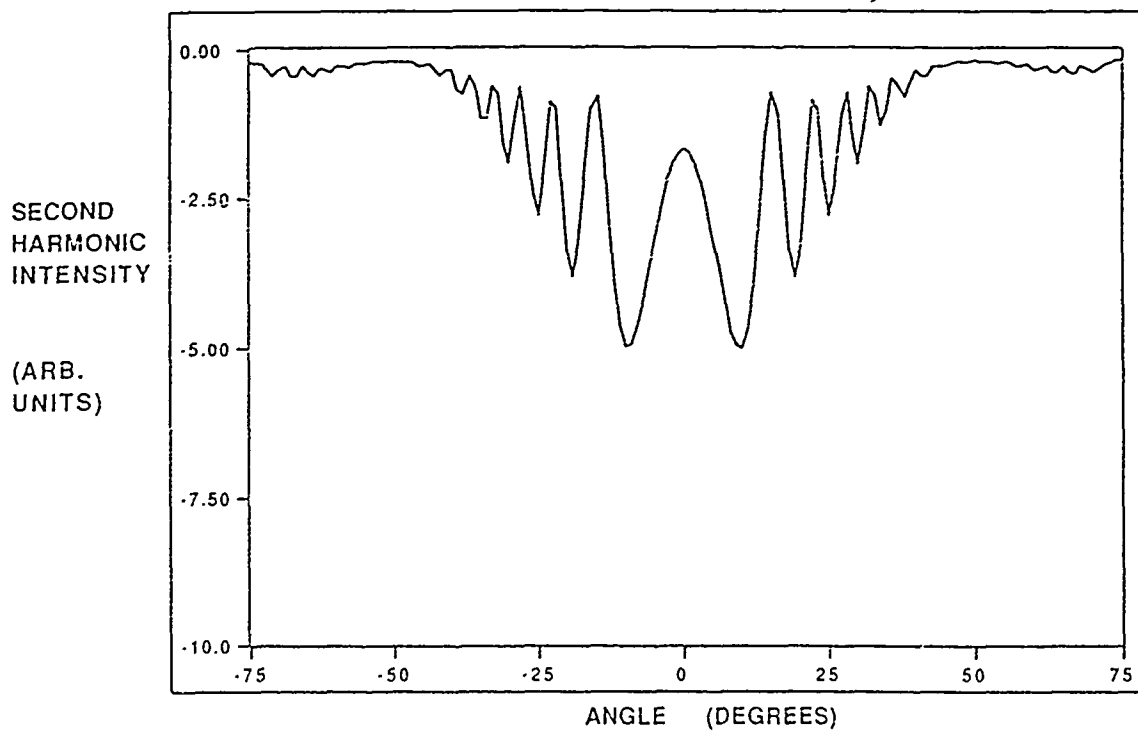


Figure 2. Maker fringe pattern of y-cut quartz.

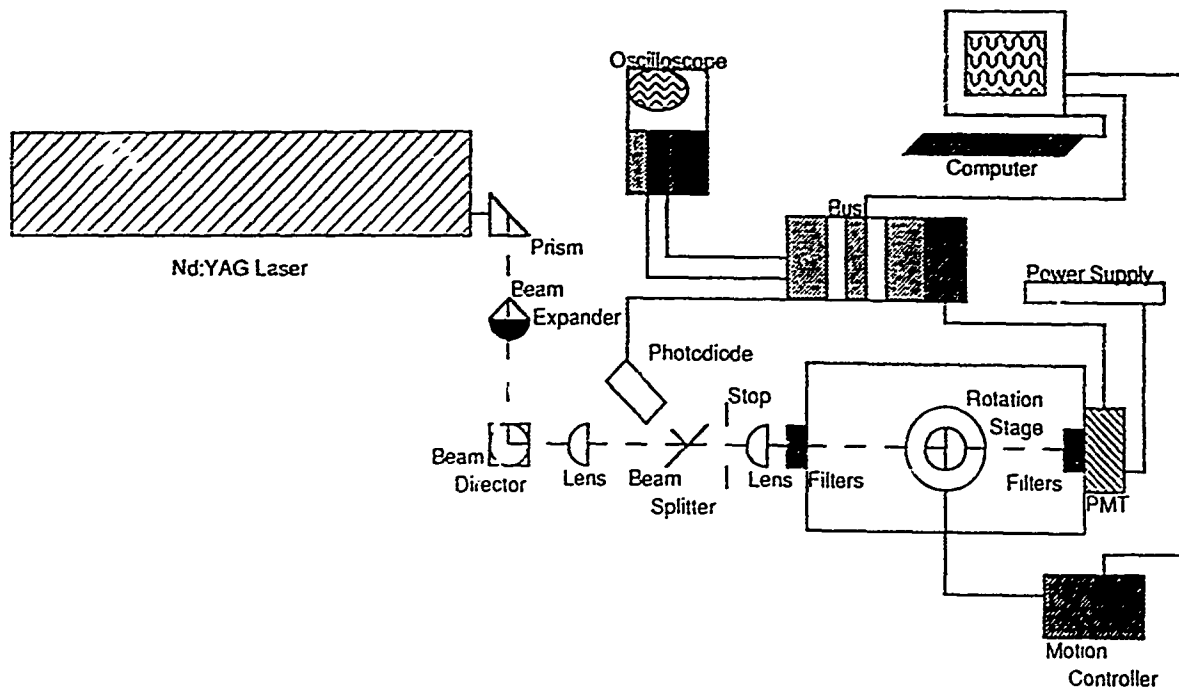


Figure 3. Experimental setup for the Maker fringe experiment.

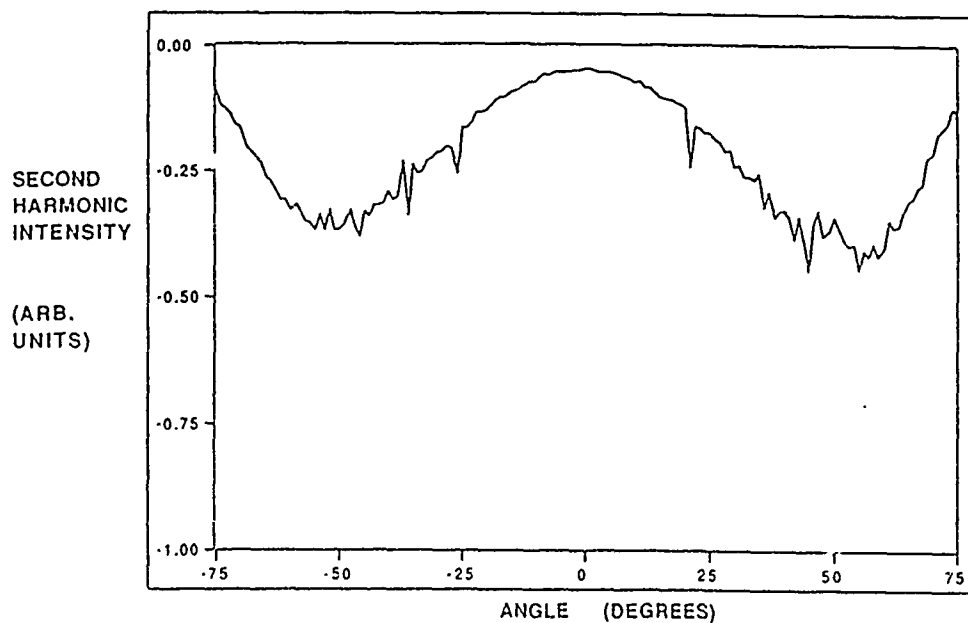


Figure 4. Maker fringe pattern for poled amorphous quartz.

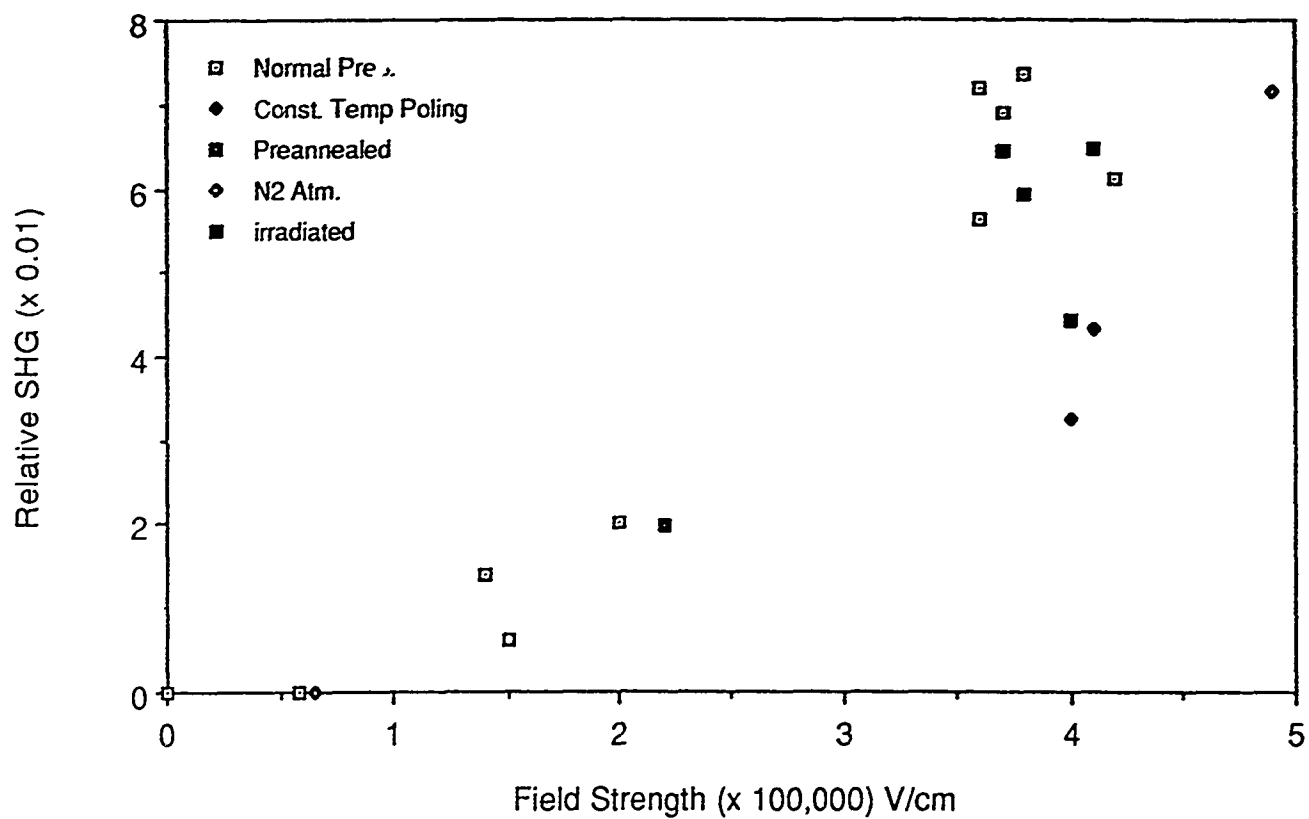


Figure 5. Dependence of SHG on the strength of the poling field. Sets of data are reported for the various conditions under which the samples were prepared (See legend).

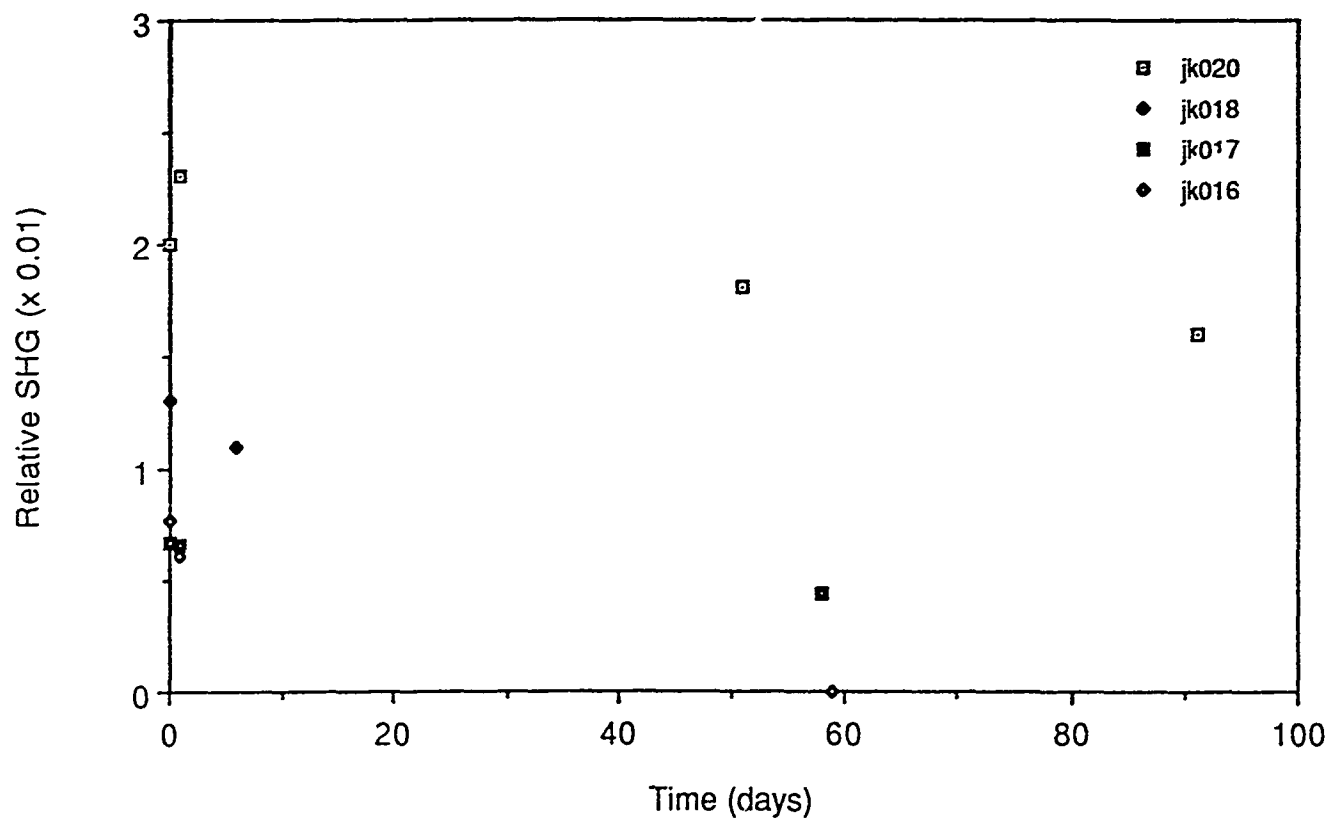


Figure 6. Relaxation of SHG effect with time. Each set of points represents the decay of SHG for a single sample.

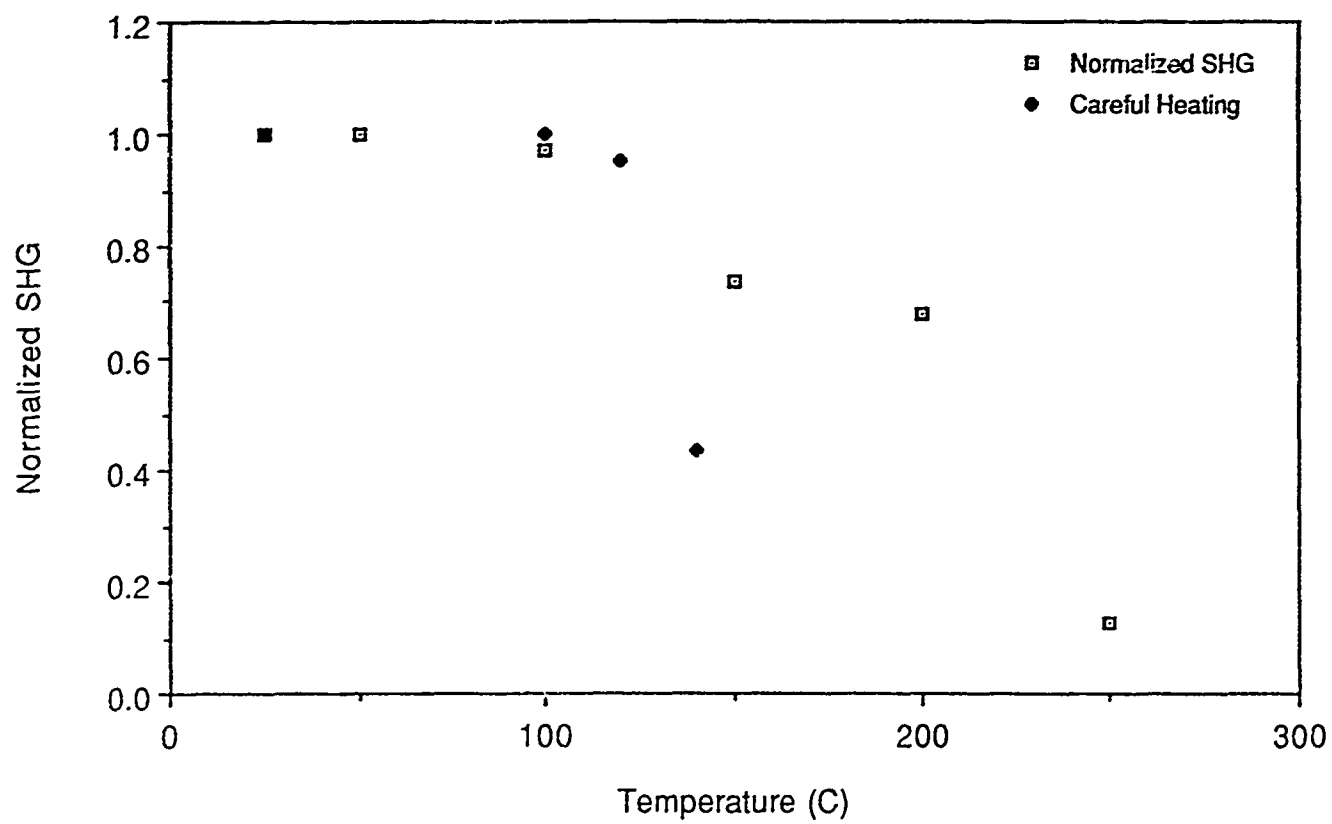


Figure 7. Dependence of SHG on subsequent heating.

| <u>Element</u> | <u>Concentration (ppb)</u> |
|----------------|----------------------------|
| Al | <5 - 50 |
| Ca | <5 - 200 |
| Cr | <1 - 26 |
| Fe | 10 - 125 |
| Mg | 5 - 70 |
| K | 5 - 170 |
| Na | 10 - 600 |
| Sb | <5* |
| As | <5* |
| Ba | <14* |
| Be | <5* |
| Bi | <10* |
| B | <100* |
| Co | <10* |
| Cu | <10* |
| Li | <1* |
| Mn | <10* |
| Mb | <5* |
| Ni | <7* |
| P | <100* |
| Ag | <50* |
| Sr | <3* |
| Th | <5* |
| Ti | <40* |
| U | <1* |
| V | <10* |
| Zn | <30* |
| Zr | <30* |
| Ga | not determined |
| Au | not determined |
| Cd | not determined |

* denotes the detection limit of the analytical technique used to measure these quantities. No impurity of these elements has been detected.

Table 1. Company specifications on concentrations of impurities in samples studied.

THREE-DIMENSIONAL FLOW STRUCTURE DEVELOPMENT AND EFFECT NEAR THE ROOT OF AN OSCILLATING WING

John D. Klinge

Abstract

Forced unsteady separated flows were studied near the root of a NACA 0015 wing undergoing sinusoidal pitch oscillations at low Reynolds numbers for various reduced frequencies and mean pitch angles. Vortex initiation, development and aerodynamic influences were characterized using hot-wire anemometry, surface static pressure measurements, and force balance measurements. The structure and development of a leading edge vortex on the wing upper surface varied temporally and spatially with different parameter combinations. Peak suction forces produced by the leading edge vortex on the wing surface were closely tied to convection and dissipation mechanisms. Forces measured on the splitter plate mounted at the root of the wing configuration were a result of the periodic vortex formation on the upper wing surface and the flow stagnation that occurred at high instantaneous angles of attack on the lower wing surface.

Introduction

Forced unsteady flow separation about an airfoil has long been studied with the intent of exploiting the associated high transient lift forces and reduced drag effects. The practical use of unsteady separated flow extends from insect flight to helicopter control. Large-scale vortical structures incorporated into unsteady flow fields and associated transient forces on lifting surfaces have been observed and quantified in two- and three-dimensional studies.

Previous investigations concerning the flow field around oscillating wing geometries have revealed strong asymmetries from wing root to wing tip in surface generated vortical structures. Vortical flows produced from unsteady separated boundary layer exhibit different initiation and growth characteristics when examined at different span locations on a finite wing. These asymmetric flow fields may yield skewed temporal load distributions on wing planforms. High performance aircraft which enter such post-stall flight regimes may be adversely affected structurally and may suffer control difficulties.

In steady state conditions, tip vortex suppression through the use of winglets on modern transport aircraft has demonstrated improved cruise performance due to the modification of the pressure distribution on wing surfaces. Under unsteady circumstances, tip vortex suppression on an oscillating flat plate has been shown to modify the three-dimensional orientation and behavior of vortical flow structures generated on the upper surface. In both steady and unsteady aerodynamics, spatially localized flows can be evaluated and controlled.

Wing/wall flow interactions can be observed in other aspects of aircraft design as well, including pylons and actuator housings. Such interactions in simple unsteady cases can extend 3-D unsteady aerodynamics knowledge from arguments of symmetry toward more reliable prediction of effects due to geometry and flight parameter variations.

Steady state investigations in the past have examined the flow field in the vicinity of a wing/body junction. Longitudinal boundary layer interaction vortices in the corner region have been measured experimentally and modeled computationally. However, little is known about flow field kinematics, spanwise variations in leading edge vortex

Previous investigations concerning the flow field around oscillating wing geometries have revealed strong asymmetries from wing root to wing tip in surface generated vortical structures. Vortical flows produced from unsteady separated boundary layer exhibit different initiation and growth characteristics when examined at different span locations on a finite wing. These asymmetric flow fields may yield skewed temporal load distributions on wing planforms. High performance aircraft which enter such post-stall flight regimes may be adversely affected structurally and may suffer control difficulties.

In steady state conditions, tip vortex suppression through the use of winglets on modern transport aircraft has demonstrated improved cruise performance due to the modification of the pressure distribution on wing surfaces. Under unsteady circumstances, tip vortex suppression on an oscillating flat plate has been shown to modify the three-dimensional orientation and behavior of vortical flow structures generated on the upper surface. In both steady and unsteady aerodynamics, spatially localized flows can be evaluated and controlled.

Wing/wall flow interactions can be observed in other aspects of aircraft design as well, including pylons and actuator housings. Such interactions in simple unsteady cases can extend 3-D unsteady aerodynamics knowledge from arguments of symmetry toward more reliable prediction of effects due to geometry and flight parameter variations.

Steady state investigations in the past have examined the flow field in the vicinity of a wing/body junction. Longitudinal boundary layer interaction vortices in the corner region have been measured experimentally and modeled computationally. However, little is known about flow field kinematics, spanwise variations in leading edge vortex

strength, and aerodynamic forces near a wing/wall junction on a planform undergoing oscillations in pitch.

The purpose of this investigation was to characterize some of the unsteady separated flow dynamics near a wing/wall junction using experimental data gathered from wing tunnel tests. A generic wing planform (Fig. 1) was driven sinusoidally in pitch about its quarterchord at a 10 degree amplitude to produce a highly complex, repeatable unsteady flow field. A fundamental understanding of root flow dynamics in this unsteady environment should lead to improvements in the design, and utilization of lifting and control surfaces for enhanced post-stall maneuvering.

Nomenclature

| | |
|-------------|--------------------------------|
| c | chord (ft) |
| f | frequency (Hz) |
| K | reduced frequency ($wc/2U$) |
| LEV | leading edge vortex |
| MPA | mean pitch angle (deg) |
| U | freestream velocity (ft/s) |
| w | angular frequency ($2\pi f$) |
| "near root" | 0.09c outboard from the root |

Results

A strain gauge instrumented splitter plate was mounted orthogonally to the oscillating wing at the root. The strain gauge measured splitter plate deflections toward and away from the wing tip. Splitter plate deflection data was ensemble averaged over 20 oscillation cycles and plotted in Fig. 2. Positive deflection voltages in these plots indicate a plate deflection toward the wing tip. A zero deflection voltage corresponds to a neutral splitter plate position, and a negative voltage indicates deflection away from the wing tip. Thorough static angle of attack testing produced data resembling

steady state, 2-D airfoil lift coefficient vs. angle of attack curves. Dynamic data magnitude variation with respect to instantaneous angle of attack bore little resemblance to data acquired in static tests. Data corresponding to deflections above the upper wing surface indicated that deflections toward the wing tip occurred, for the most part, in the first half of the oscillation cycle. The splitter plate then tended to arrive at the neutral position some time near the middle of the cycle and pass through or dwell there before deflecting towards the wing tip again. Data corresponding to the deflections below the lower wing surface show deflection histories opposite in phase to matching oscillation conditions on the upper surface. These data indicated a splitter plate deflection away from the wing tip in the first half of the cycle, followed by a deflection back to the neutral position in the second half.

At each reduced frequency (0.10, 0.15, and 0.25), average deflection magnitudes corresponding to the splitter plate mounted above the upper surface increased with mean pitch angle. In addition, for matching oscillation parameters, the negative deflection magnitudes under the lower wing surface were similar to positive deflection magnitudes above the upper wing surface.

For mean pitch angles of +10 and +20 degrees, clear trends were apparent in the splitter plate deflection data. As K increased, the maximum deflection toward the wing tip occurred later in the cycle. As the mean pitch angle increased, the maximum deflection toward the wing tip occurred earlier in the cycle. At a mean pitch angle of 0 degrees, appreciable variations in deflection peak time did not occur with reduced frequency.

Fifteen static pressure transducers were mounted inside the wing cavity, and close coupled to the surface through pressure ports cut into the wing. These ports extended

from the leading edge of the wing to the trailing edge in a straight, chordwise line. The pressure ports could be effectively moved to any span position on the wing by using tip extensions and a movable splitter plate. Data collected from the pressure transducers were taken two oscillation cycles at a time and ensemble averaged over twenty consecutive runs. Pressure coefficient data taken at a reduced frequency of 0.25 are plotted in Fig. 3. This plot shows two averaged cycles of data taken at various chord positions from the leading edge (bottom signal) to the trailing edge (top signal). Absolute signal magnitudes are not quantified in Fig. 3. However, local suction peaks are designated for each signal by the circular symbols. The suction peak time and magnitude were extracted from plots like this one and graphed in Fig. 4. The suction peak convection history was similar but phase shifted for mean pitch angles of +10 and +20 degrees, with initiation occurring closer to the end of the oscillation cycle for the +10 degree case (Fig. 4, top). For all cases shown in Fig. 4, the peaks were recorded first at approximately 5% chord. Suction peaks then appeared to spread both upstream and downstream from the 5% chord position. These peaks initially spread from 0% chord upstream to nearly 17% chord downstream for the +20 degree mean pitch angle case, and from 0% upstream to nearly 7% downstream for the +10 mean pitch angle case. After the initial spread, suction peaks convected downstream over the surface of the wing to the trailing edge before disappearing. Suction peaks recorded near the root appeared initially to lag behind the peaks recorded at the center span in terms of chord position. Just prior to 50% cycle time, however, root peaks caught up to center span suction peaks and completed convection to the trailing edge earlier in the cycle. This phenomenon was more pronounced at the higher mean pitch angle.

The bottom graph in Fig. 4 shows suction peak magnitudes that correspond to the convection histories plotted directly above. The average magnitude of the suction peak seen at $MPA = +20$ degrees was higher than that seen at $MPA = +10$ degrees.

However, the range of the +10 degree mean angle data in Fig. 4 (bottom) exceeded that of the +20 degree mean angle data. Data taken near the root was always greater in magnitude in the early peak convection stages, and lesser in magnitude near the end of peak convection.

A single element low-interference hot-wire anemometer probe was mounted to the surface of the wing at various locations. Flow velocity data was collected and ensemble averaged over 20 oscillation cycles. Velocity profiles corresponding to the time and location of suction peak occurrence at center span and near the root are presented in Fig. 5. In these plots, the change in velocity with respect to the vertical distance from the wing surface is indicated. It should be noted that these profiles are not true velocity profiles because flow direction was not measured. However, a high percentage of the flow near a wing is likely to be parallel to the surface. Reverse flow near the plate was not identified.

The top graph in Fig. 5 shows the cross-section of a vortical structure at center span and near the root, at the 33.3% chord location during a local suction peak. Most velocity change at center span occurs between 10% and 20% chord above the surface. A peak velocity of 140% freestream occurs 25% chord off the wing surface. Near the root, all velocity features are located closer to the wing surface. The bottom graph in Fig. 5 shows center span and near root span velocity profiles at the 66.7% chord location during local suction peaks. These profiles presumably show the same vortical structure as discussed above, later in the oscillation cycle and farther downstream. At center span, characteristics seen upstream (in the top graph) have largely dissipated. The velocity gradient is less pronounced, and the vertical distances between measured maximum and minimum velocities have grown. Near the root (Fig. 5, bottom), however, the structure profile has retained characteristics noted upstream (Fig. 5, top).

The sharp velocity gradient occurs over a greater distance and higher off the wing surface than at the upstream/near-root location and center span location.

Fig. 6 shows velocity profiles taken normal to the splitter plate, 50% chord above the wing surface. These profiles are perpendicular to those shown in the bottom graph in Fig. 5, occupying the same plane cutting through the 66.7% chord position along the span of the wing. All profiles were measured well outside the periphery of the leading edge vortex passing over the wing upper surface. Throughout the oscillation cycle, the splitter plate velocity profiles maintained a pseudo-steady state appearance. A boundary layer measuring 0.3 or 0.4 inches in thickness was the only apparent feature. In Fig. 6, the profile corresponding to 14% cycle time was taken during a local suction peak on the upper wing surface at the same chord location. The other profile is opposite in phase with respect to the oscillation cycle. Maximum velocities observed in these profiles are greater in the early part of the cycle, while boundary layer thickness is greater in the later part of the cycle. These differences varied gradually throughout the oscillation cycle, peaking at those cycle times corresponding to Fig. 6.

Discussion

Splitter plate deflection histories revealed trends in peak deflection magnitudes across the parameter range that were identical to those seen for leading edge vortex initiation in oscillating flat plate experiments performed previously by other investigators, and in oscillating wing flow visualization tests performed previously by this writer. Peak splitter plate deflection over the upper wing surface toward the wing tip occurred earlier in the oscillation cycle for increased mean pitch angles, and later in the cycle for increased reduced frequencies. Peak splitter plate deflection seen in this investigation occurred simultaneously with leading edge vortex development on the wing upper surface visualized in preliminary tests. Splitter plate deflection over the upper surface toward the wing tip can be attributed to leading edge vortex presence. High local

velocities generally associated with the energetic leading edge vortex and a high instantaneous angle of attack were likely to have caused temporal variations in static pressure on the plate surface.

On the wing-side of the splitter plate during maximum pitch up (0% cycle time), flow stagnation presumably occurred on the underside of the wing creating local high pressures on the splitter plate surface. On the side of the splitter plate opposite to the wing, the relatively unperturbed freestream flow exerted the usual amount of static force on the splitter plate. The net difference between these static pressures on the two splitter plate surfaces resulted in a deflection away from the wing tip near maximum pitch up.

The deflection histories of the splitter plate above the wing upper surface and below the wing lower surface were remarkably similar for matching oscillation conditions. While the splitter plate above the wing upper surface was deflecting toward the wing tip, the splitter plate below the wing lower surface was deflecting away from the wing tip at roughly the same magnitude. No mechanical coupling between these phenomena was possible; therefore, the overall periodic torque on the splitter plate configuration was a result of fluid mechanisms on both surfaces of the wing acting simultaneously, with equal force, in opposite directions. The overall effect of variation in static pressure on the splitter plate above the upper wing surface due to LEV influence was the same as that below the lower wing surface due to flow stagnation.

Suction peak convection data shown in Fig. 4 (top) also correspond well with leading edge vortex convection histories generated from flow visualization in previous experiments. These suction peaks are therefore presumed to be due to high localized velocities associated with the leading edge vortex. As the suction peak began to

convect downstream over the upper surface of the wing, root positions lagged upstream. When the suction peak neared the 50% chord position, the root segment overtook the center span segment. Convection to the trailing edge occurred earlier in the cycle for the root segment.

Negative pressure coefficients (localized to the leading edge vortex position as discussed above) in the bottom of Fig. 4 indicate that the greatest local suction forces were exerted on the wing during LEV initiation and development. Dissipation mechanisms presumably governed observable LEV growth and strength loss. Since the visualized LEV remained attached to the upper surface of the wing from initiation to shed time in all previous experiments, LEV dissipation undoubtedly accounts for the dominant trend in the suction peak magnitude histories plotted in Fig. 4 (bottom). However, a systematic relationship between LEV convection rate (Fig. 4, top) and some of the more subtle trends in suction peak magnitude variation (Fig. 4, bottom) are apparent when comparing pressure measurements taken near the root with those taken at center span. For example, in Fig. 4 (top) at any given mean pitch angle, when the slope of the center span curve exceeds that of the "near-root span" curve, corresponding suction peak magnitudes (Fig. 4, bottom) are lower at the center span compared to near the root. Higher leading edge vortex convection rates resulted in smaller local suction forces. While not surprising under ideal conditions, this relationship played a primary role in suction peak magnitude variations seen locally on the surface of the oscillating wing.

Common features in LEV velocity profiles, such as steep velocity gradients and maximum/minimum velocities, were always located closer to the wing surface near the root than they were at center span. LEV profiles near the root (Fig. 5) consisted of higher maximal velocities and sharper velocity gradients, indicating stronger vortices

than those seen at center span. At higher reduced frequencies ($K = 0.5, 0.75$) the opposite was true: LEV profiles not shown indicated stronger vortices at center span than those seen near the root.

Velocity profiles depicting splitter plate surface flow (Fig. 6) did not indicate the presence of consolidated vorticity at any point in the cycle for any of the reduced frequencies tested ($K = 0.25, 0.5, 0.75$). All profiles collected exhibited only a boundary layer velocity distribution. The fluctuating boundary layer thickness was continuously less than twice the theoretical Blasius solution for two-dimensional, steady-state, incompressible flow under similar circumstances.

Conclusions

Near the root of the oscillating wing, the leading edge vortex dissipated into the splitter plate boundary layer, turning neither upstream, downstream, or up the surface of the splitter plate. The leading edge vortex was compressed slightly on the wing surface near the root compared to center span.

Suction peak convection histories correlated well with previously established leading edge vortex convection. Convection near the root and at center span was different. The initially slow LEV root segment eventually caught and passed the center span segment before shedding into the wake of the wing. Higher convection rates resulted in smaller suction forces recorded at the wing surface directly under the LEV. LEV dissipation and convection governed local pressure coefficient trends on the wing surface.

Unsteady aerodynamic forces on the splitter plate were governed by leading edge vortex presence on the wing upper surface, and flow stagnation on the lower wing surface. These two phenomena and the associated in-phase velocity changes in the

splitter plate surface flow produced a periodic torque on the splitter plate configuration -- forcing it to deflect toward the wing tip above the wing upper surface while simultaneously deflecting it away from the wing tip below the lower wing surface.

Some of the macroscopic features of three-dimensional forced unsteady separated flow were successfully detected and outlined by these experiments under limited and simplified conditions. Similar methods could be employed not only to study more complex aerodynamic variations in model geometry, motion, and flight parameters, but also to conduct well-controlled, detailed, time-intensive investigations into the intricate nature of unsteady flow fields. The results of this investigation provide a simple bench-mark for expanded and more detailed studies that will lead to more efficient control and exploitation of unsteady separated flow fields under applied circumstances.

List of Figures

1. Oscillating wing configuration. Circular splitter plate, at left, is mounted orthogonally to NACA 0015 wing at the root. Arrow indicates freestream velocity vector. Measurements are span = 12", chord = 6", plate diameter = 12".
2. Splitter plate deflection histories. Conditions are $K=0.10$ (top), $K=0.15$ (middle), $K=0.25$ (bottom). MPA = +10a refers to deflections above the wing upper surface, and MPA = +10b refers to deflections below the lower wing surface. Beginning of the oscillation cycle (0% cycle time) corresponds to maximum pitch up, 50% cycle time corresponds to maximum pitch down.
3. Two oscillation cycles of static pressure transducer signals with local suction peaks designated by circular symbols. Signals were collected from pressure ports located chordwise on the surface of the wing from the leading edge (bottom signal) to

90% chord (top signal). Horizontal axis corresponds to time. Conditions are $K=0.25$, $MPA=10$ deg., upper wing surface, center span.

4. Suction peak convection history (top) and corresponding suction peak magnitudes (bottom). "C.S." refers to center span, "N.R." refers to near the root. Conditions are $K=0.25$, upper wing surface.

5. Leading edge vortex velocity profiles. Conditions are $K=0.25$, $MPA=10$ deg. (all); chord location = 33.3% (top); chord location = 66.7% (bottom).

6. Splitter plate surface flow velocity profiles. Conditions are $K=0.25$, $MPA=10$ deg.

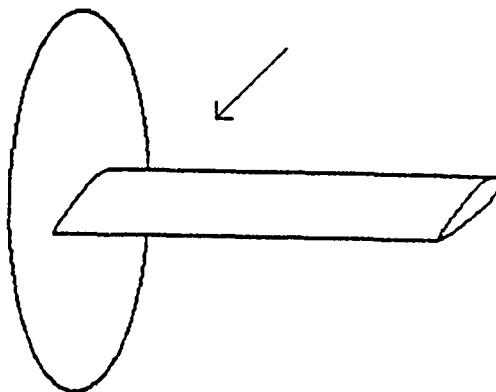


Fig. 1

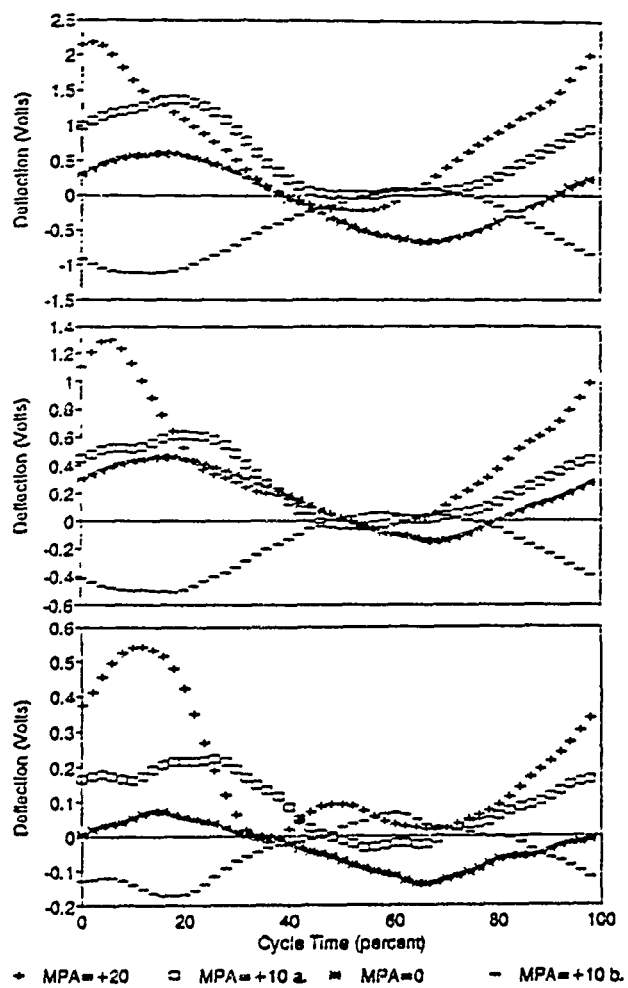


Fig. 2

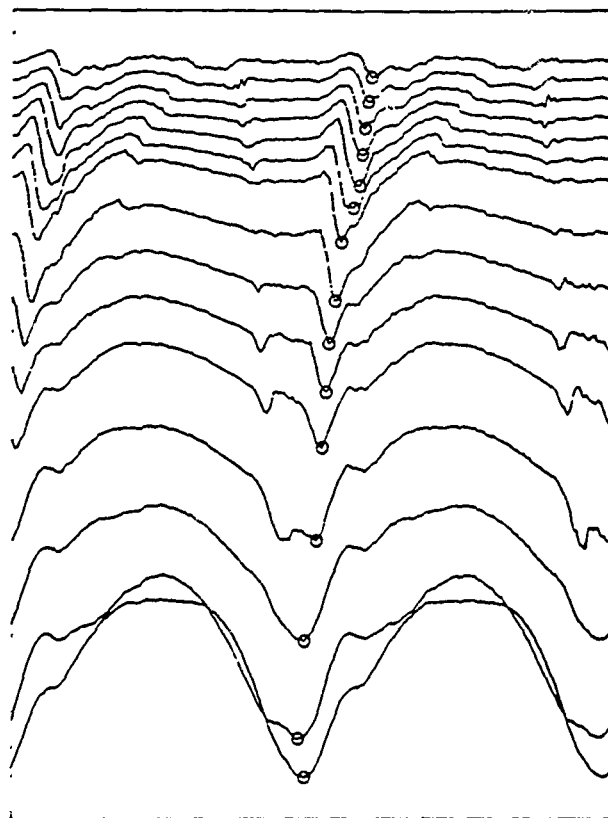


Fig. 3

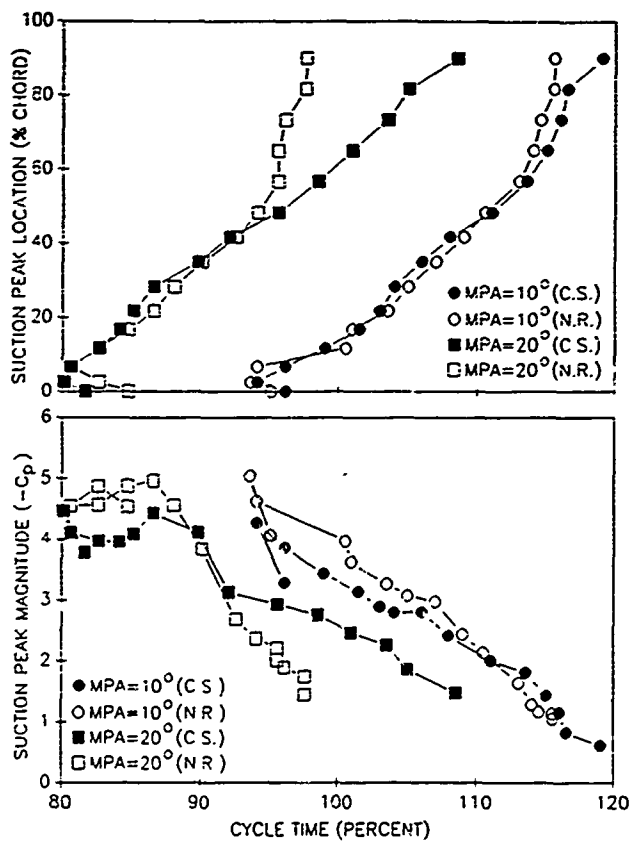


Fig. 4

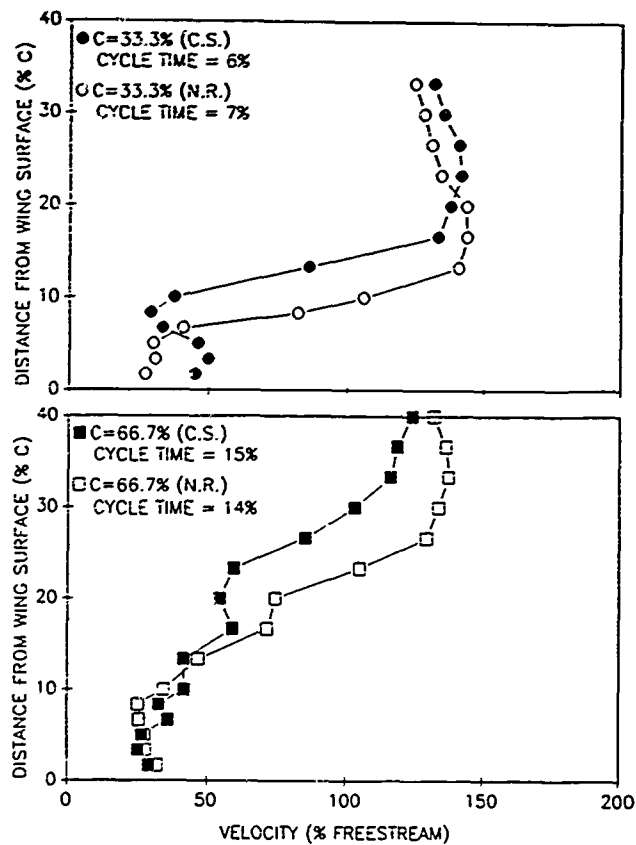


Fig. 5

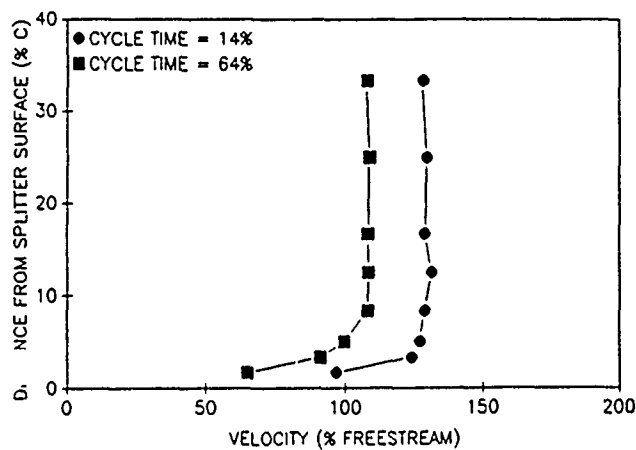


Fig. 6

AN INVESTIGATION OF ACTIVELY CONTROLLED STATIC AND DYNAMIC AIRFOILS

Julie A. Lovato, PhD Candidate

Washington State University

ABSTRACT

This research investigation evaluates the use of internal tangential pulsed air forcing as a means of steady and unsteady separation control on a NACA-0015 airfoil. Control frequencies are related to the natural vortex passage frequency associated with a separating static airfoil flow. Results show promise in reducing the separation region on both the static and dynamic airfoil flows. The static airfoil shows two distinct phenomena caused by the pulsed air forcing, a leading edge reattachment region and a clockwise vortex downstream of the leading edge. Both phenomena appear to be frequency dependent. The dynamic stall vortex on an unsteady airfoil is eliminated or reduced for certain forcing frequencies. The control effects involve interactions between pulse-generated vortices and the dynamic stall vortex generated by the unsteady flow.

INTRODUCTION

The large transient aerodynamic forces generated by unsteady separation hold promise in augmenting aircraft performance. Increases in lift coefficients of up to five times static values have been obtained from

airfoils undergoing unsteady motions.¹ Taking advantage of such forces is limited due to their fleeting occurrence on the airfoil during its unsteady pitching motion. The dramatic increases in lift are followed by the growth and detachment of a large leading edge vortex which affects airfoil performance and accompanies catastrophic stall.

The relationship between the leading edge, or dynamic stall, vortex and the accompanying airfoil flow characteristics and aerodynamic forces has been well documented by past researchers.²⁻⁴ In particular, it has been determined that the characteristic lift increase in these unsteady flows occurs prior to the rapid growth of the vortex.⁵ Thus it is reasonable to focus on the flow preceding vortex formation when developing control strategies for maintaining enhanced lift over longer periods. The research effort discussed here draws a qualitative analogy between flow over static airfoils and flow over airfoils pitched at constant rates. Except for the fact that the unsteady flow remains attached for higher angles of attack, the two flows are qualitatively similar up to the formation of the unsteady dynamic stall vortex. As the angle of attack increases, the flow transitions from a quasi-steady, symmetric flow to a separating flow with an increase in the strength of the trailing edge vortices. As the separation zone grows a shear layer interaction between the separating boundary layer and the separated flow produces a number of small clockwise vortices. This continues until, on the steady airfoil, the entire upper surface has separated and the airfoil is stalled. On the dynamic airfoil the dynamic stall vortex grows and eventually departs from the airfoil surface.

It is the similarities between the steady and unsteady flows prior to dynamic stall vortex formation which lead to the control approach explored in this study. According to Gad-el-Hak and Bushnell⁶, the separating boundary layer on a static airfoil can be considered a boundary layer

transitioning to a free shear layer. Assuming that the separating boundary layer on an unsteady airfoil is also similar to a free shear layer, then control mechanisms which have been successful in affecting free shear layer growth and that will control static airfoil separation can be applied to the unsteady case as well. These control methods revolve around the manipulation of instabilities in the separating flow. The instabilities, known as Kelvin-Helmholtz instabilities in a laminar flow, create small perturbations in the flow. The two-dimensional waves grow exponentially with downstream distance and roll up into vortices.⁷ It has been proven that forcing at the generating frequency of these large scale vortices or at the accompanying subharmonics significantly affects shear layer behavior.⁸

The research discussed in this report uses the method of tangential pulsed air to actively control the flow over static and dynamic airfoils. The use of single pulse injection control has been applied successfully to an unsteady flow. In their study of an oscillating airfoil, Robinson and Luttges⁹ found that injecting a single pulse of air perpendicular to the airfoil surface could create a vortex similarly oriented to the dynamic stall vortex generated by the airfoil unsteady motion. Depending upon their phase relationship, the two vortices interacted and at times successfully enhanced the unsteady flow control. The control method applied in this report uses the same idea, but instead of a single pulse inputs the air tangentially at frequencies related to the natural flow itself.

OBJECTIVE OF THE RESEARCH EFFORT

The objective of this research project was to assess feasibility of low frequency flow control of static and dynamic two-dimensional airfoils. Both types of flow were forced using low frequency internal pulsed air

coupled with natural static flow frequencies. Flow visualization was used to analyze the qualitative flow behavior changes with respect to forcing frequency.

EXPERIMENTAL PROCEDURES

The experiments were conducted in the Frank J. Seiler Research Laboratory's open return, low speed wind tunnel at the U. S. Air Force Academy. The wind tunnel has a 0.91 m x 0.91 m test section designed for use with flow visualization and flow sensor measurements. All experiments were conducted at a freestream velocity of 3.0 m/s, with turbulence intensity levels below 0.05%. A NACA-0015 airfoil with a 61 cm span and a 15 cm chord was pitched about the quarter chord axis at constant rates from 0° to 50°.

The flow was visualized by placing a smoke wire 30 cm upstream of the airfoil leading edge. Motor oil was applied to the horizontal 0.127 mm tungsten wire, leaving fine droplets almost uniformly along the wire. A current applied to the wire evaporated the oil, producing fine streaklines across the test section. This method resulted in a clear smoke outline of the separation region and the dynamic stall vortex as it formed on the airfoil surface. The streaklines were illuminated using EG&G Strobebrite strobe lights with an average 7 μ s flash duration. A trigger box allowed synchronization of a FE-II Nikon camera with the strobe lights, smoke wire, and the pitching mechanism to yield photographs at a given instant in the flow. A schematic of the flow visualization design is shown in Figure 1.

A key element in this study was the determination of the shear layer structure behavior in both the controlled and natural flows. Before active control could be applied to the flow, the fundamental frequency of the large-scale shear layer structures had to be determined. The fundamental

vortex passage frequency is the frequency at which the large-scale shear layer structures are generated in the separating boundary layer. The frequency was obtained by placing a hot-film near the shear layer to sense the vortices as they were created, but not so close that it experienced the turbulence of the shear layer itself. An example of the corresponding signal is shown in Figure 2(a). The signal has the pattern of a doppler burst, which is typical of naturally occurring random phenomena.

For this study, the edge of the separated shear layer was determined from a combination of flow visualization observations and the presence of a clear doppler signal on the oscilloscope. The flow was then forced at an amplitude corresponding to 0.05 volts rms on the Fluke Model 8050A digital multimeter. The frequency was set to that obtained from the doppler-type signal, and then adjusted until the characteristic trace of a flow forced at the fundamental vortex passage frequency was obtained. An example of the oscilloscope trace of a flow forced at its fundamental frequency is shown in Figure 2(b). This frequency was then recorded as the fundamental vortex passage frequency. For the flow over a static airfoil, this frequency is dependent upon both free stream velocity and angle of attack. Katz¹⁰ determined that the fundamental frequency found in the wake of a NACA-0012 airfoil could be represented by a Strouhal number of 0.18 - 0.19, with the Strouhal number being calculated using the equation $St = f(\sin\alpha)c/U_\infty$. In this case, f is the vortex passage frequency, α is the angle of attack, c is the chordlength, and U_∞ is the freestream velocity. Measuring the fundamental frequencies nearer to the leading edge in this experiment resulted in the same relationship, with the Strouhal number equal to four times that of the wake, or in the range 0.70 - 0.75.

The acoustic control used to determine the fundamental static flow frequencies was supplied externally to the airfoil by a speaker. A 220 Watt woofer, mounted as shown in Figure 3, was modulated with a sine wave generator, at frequencies between 20 HZ and 300 HZ. The signal was amplified by a Crown PS-400 Dual Channel power amplifier, with the sound being transmitted through a ducting system to a location over the leading edge region of the airfoil. Two inch diameter PVC pipe was used to transmit the acoustic control, supplying the signal with minimum distortion. Forcing amplitude, measured by a microphone placed at the pipe exit, was kept at a constant 0.05 rms volts.

The primary forcing using internal tangential-pulsed air control was accomplished through three downstream facing slots located on the upper surface of the airfoil. The slots were 46 cm in the spanwise direction and 1.6 mm wide. They were mounted flush with the airfoil surface in order to minimize airfoil distortion. As shown in Figure 4, the system employed six solenoid valves capable of pulsing between zero and 40 HZ. The valves were controlled through a timing circuit by a square wave generator. Compressed air was supplied at 30 psi through the valves using tygon tubing.

RESULTS

The first part of the research effort involved a study of active control effects on the flow over a static airfoil. Figure 5 shows the flow forced at frequencies corresponding to the fundamental and subharmonics of the large scale structure passage frequency. Each subsequent subharmonic referred to in this report is equal to the previous subharmonic divided by two. Forcing at frequencies greater than 40 HZ was not possible due to the design constraints of the pulsed air system. As such, certain flow

frequencies could not be applied. The angles of attack of 10° , 15° , 20° , and 25° were analyzed. For a 10° angle of attack, shown in column one, forcing at the second subharmonic results in significant reattachment near the leading 10% of the airfoil surface. At approximately $0.4c$, corresponding to the location of the third pulsing slot, the separation region grows rapidly as the flow detaches from the airfoil. This increase in separation from the unforced case could be a result of enhanced shear layer structure pairing. Inducing mixing layer pairing by forcing at subharmonics of the fundamental vortex passage frequency has been shown to enhance mixing layer growth.⁸ Forcing at the third subharmonic results in a slight overall reduction in separation.

Forcing at the first and second subharmonics for an attack angle of 15° , shown in column two, results in the amplified growth of large scale structures in the separating boundary layer. As at the 10° attack angle, there is a significant initial reduction in separation, followed by a separation region that is greater than the natural case. Forcing at the third subharmonic has only minimal effect on the flow. Enhancement of large scale shear layer structure development is still apparent at an attack angle of 20° forced at the first subharmonic, as seen in column three. The structures are clearly of a clockwise nature, and appear more organized due to their greater boundary definition. As at the previous attack angles, there is significant reattachment of the flow over the first half of the airfoil when the flow is forced at the lower subharmonics.

In column four, the attack angle of 25° forced at the fundamental frequency shows a disruption in the shear layer structure growth that is apparent in the natural case. This could be the result of pairing inhibition, a known result of forcing at the fundamental frequency.⁸ The

small reattachment region near the airfoil leading edge is still present in the flow forced at the first subharmonic, and is followed by a large clockwise vortex which covers approximately 30% of the chord. At the lower subharmonics, the small leading edge reattachment region has been replaced by a loose clockwise structure encompassing the entire airfoil. This results in a reattachment of the flow near the airfoil trailing edge.

A more comprehensive study concerning the dependence of static flow behavior on forcing frequency was conducted for an angle of attack of 25° . At the higher frequencies, shown in Figure 6(a) at 36 Hz and Figure 6(b) at 31 Hz, there are only subtle effects on the flow. At 36 Hz, the large scale structures in the separating boundary layer do not appear to pair as quickly and grow with downstream distance as they did in the natural flow. Instead, they are shed off the airfoil leading edge and only loosely combine as they pass along the chord. At 31 Hz, there is an indication of a short reattachment region covering approximately the leading 5% of the chord. The overall separation region has been reduced slightly, indicated by the movement of the smoke streaklines closer to the airfoil surface.

The result of forcing at the mid-range frequencies is shown in Figures 6(c) – 6(e). At these frequencies, there are two distinct areas of interest on the airfoil surface. The first is the short reattachment region near the airfoil leading edge, which grows both in the chordwise direction and in height with respect to the airfoil. There is also a loosely defined, larger vortex just downstream of this region. This vortex grows as the forcing frequency is decreased, resulting in further reduction in the airfoil separation region. At a forcing frequency of 11 Hz, shown in Figure 6(f), the leading edge reattachment region has been replaced by a large leading edge vortex. There are also indications of a large structure

just downstream of the leading edge vortex, possibly the remnants of the vortex present at the greater forcing frequencies. As the forcing frequency is reduced to 6 Hz, shown in Figure 6(g), the leading edge vortex grows in the chordwise direction until it encompasses the entire airfoil surface. The presence of the leading edge vortex results in a marked reattachment of the flow over the unforced case.

The effect of internal tangential pulsing on a dynamic flow at a nondimensional frequency of $\alpha^+ = 0.05$ depends on forcing frequency, as shown in Figure 7. At an attack angle of 20° , shown in column one, forcing frequencies in the range below 11 Hz slightly reduce the separated, turbulent region over the airfoil. At higher frequencies, the reattachment region near the leading edge that was present in the static cases is again evident. Downstream of this area, the boundary layer proceeds to separate similarly to the unforced case.

The controlled flow at an angle of attack corresponding to 25° , shown in column two, exhibits marked differences from the natural flow. At the highest forcing frequency of 22 Hz, a small leading edge vortex is present, followed by a turbulent region comparable in size to the dynamic stall vortex. There is, however, no clear indication of a stall vortex developing on the airfoil surface. As the forcing frequencies are reduced further, the leading edge vortex appears to grow and interact with the turbulent region downstream. An exact interpretation of interaction cannot be deduced from the still photographs. The result is an overall reduction in separation, and an apparent delay in growth of the dynamic stall vortex.

Shown in column three, the flow at an attack angle of 30° also shows a marked difference under forced conditions. At the highest frequencies, three separate regions are evident. A small vortical region near the

leading edge is followed by a larger structure, which appears to be interacting with another structure on a scale with the chord. Comparing this with a similarly forced static flow at 30° , the two small structures are those caused by the pulsing action, while the largest structure downstream is actually the dynamic stall vortex generated by the unsteady motion of the airfoil. The distinction between the pulsed-generated structures and the dynamic stall vortex is still clear at the lower forcing frequency of 18 Hz. At 11 Hz, however, it is difficult to discern exact boundaries between the two types of structures. In addition, the overall height of the separation region is slightly reduced. It should be noted that, in the static case at this attack angle and frequency, the secondary vortex has disappeared and only the leading edge vortex remains. At the lower frequencies of 9 Hz and 6 Hz, the distinction between the two types of structures has again become evident.

CONCLUSIONS

The internal tangential pulsing control method proved successful in forcing both steady and unsteady airfoil separation. Active control of the static airfoil resulted in the manipulation of shear layer vortex growth and the reduction of the separation region at higher attack angles. The development of two distinct regions was found at the highest attack angle studied, 25° . The flow over the airfoil was characterized by a small reattachment region near the airfoil leading edge, followed by a larger clockwise vortex which resulted in significant flow reattachment. A study at higher attack angles will be necessary to determine both the relationship between these two flow structures and whether or not they are a condition related to forcing at specific subharmonics.

Pulsing air at static fundamental and subharmonic frequencies into a dynamic flow resulted in the apparent attenuation of the dynamic stall vortex for certain cases. Depending upon forcing frequency, the overall height of the separation region was reduced, effectively extending the flow reattachment for a greater period. This might lead to a possible lengthening of the pitch cycle percentage under which an airfoil experiences enhanced lift.

RECOMMENDATIONS

Two key areas must be addressed before additional research is done in this area. First, a continuous film must be taken of the static and dynamic flows under forced conditions to determine both the relationships between the two control-related flow regions evident on the static airfoil and the interaction of these regions with the dynamic stall vortex induced by the unsteady motion. This could be augmented by a study evaluating control through each slot individually. Second, a surface pressure study must be conducted to determine aerodynamic force enhancements associated with the observed flow alterations occurring under controlled conditions.

ACKNOWLEDGEMENTS

The author would like to thank the Air Force Office of Scientific Research (Air Force Systems Command) and especially the Frank J. Seiler Research Laboratory for sponsoring this project. Special thanks are given to Captains Scott Schreck and David Bunker and Dr. Timothy Troutt for their valuable assistance during the course of this effort. Mr. Bob Hatfield and SSgt Young Paek deserve many thanks for their time and efforts in the numerous experimental systems constructed for the research project.

REFERENCES

1. Carr, L. W., McAlister, K. W., and McCroskey, W. J. (1977) "Analysis of the Development of Dynamic Stall Based Upon Oscillating Airfoil Experiments," NASA TN-8382.
2. Albertson, J. A., Troutt, T. R., and Kedzie, C. R. (1988) "Unsteady Aerodynamic Forces at Low Airfoil Pitching Rates," AIAA Paper No. 88-2579.
3. Carr, L. W. (1988) "Progress in Analysis and Production of Dynamic Stall," J. Aircraft, Vol 25, No. 1, pp 6-17.
4. Walker, J. M. and Chou, D. C. (1987) "Forced Unsteady Vortex Flows Driven by Pitching Airfoils," AIAA Paper No. 87-1331.
5. Albertson, J. A., Troutt, T. R., and Kedzie, C. R. (1988) "Unsteady Aerodynamic Forces at Low Airfoil Pitching Rates," AIAA Paper No. 88-2579.
6. Gad-el-Hak, M. and Bushnell, D. (1991) "Status and Outlook of Flow Separation Control," AIAA Paper No. 91-0037.
7. Ho, C. M. and Huerre, P. (1984) "Perturbed Free Shear Layers," Ann. Rev. Fluid Mech., Vol. 16, pp 365-424.
8. Oster, D. and Wygnanski, I. (1982) "The Forced Mixing Layer Between Parallel Streams," J. Fluid Mech., Vol. 123, pp 91-130.
9. Robinson, M. C. and Luttges, M. W. (1986) "Vortices Produced by Air Pulse Injection from the Surface of an Oscillating Airfoil," AIAA Paper No. 86-0118.
10. Katz, J. (1981) "A Discrete Vortex Method for the Non-Steady Separated Flow Over an Airfoil," J. Fluid Mech., Vol. 102, pp 315-328.

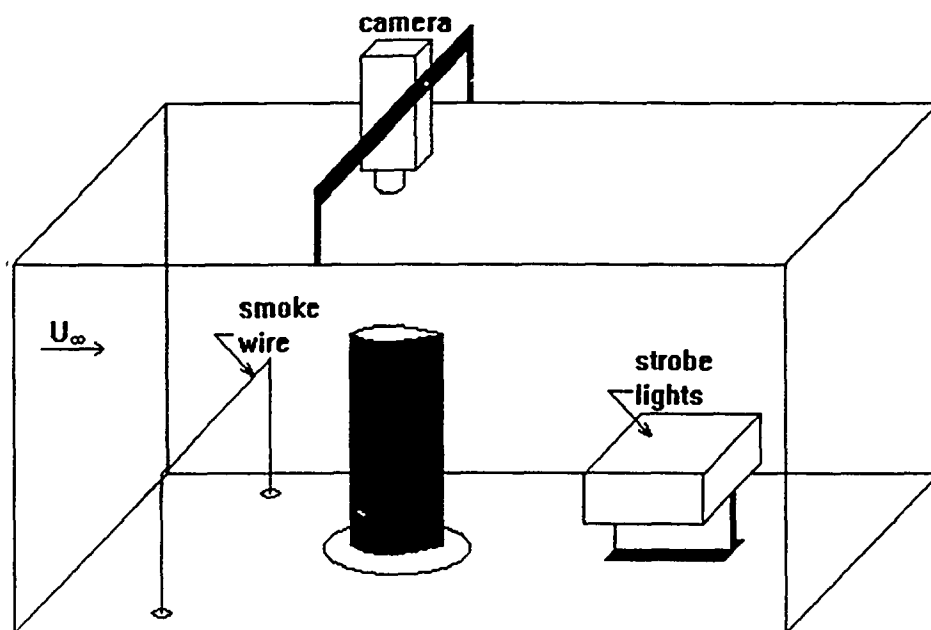
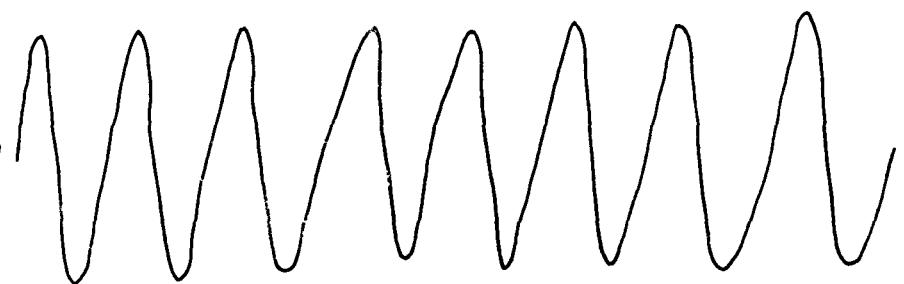


Figure 1. Flow Visualization Configuration.



(a) Sample time trace exhibiting the fundamental vortex passage frequency.



(b) Sample time trace of a flow forced at the fundamental vortex passage frequency.

Figure 2. Sample time traces of an unforced vortex passage signature and a flow forced at the fundamental vortex passage frequency.

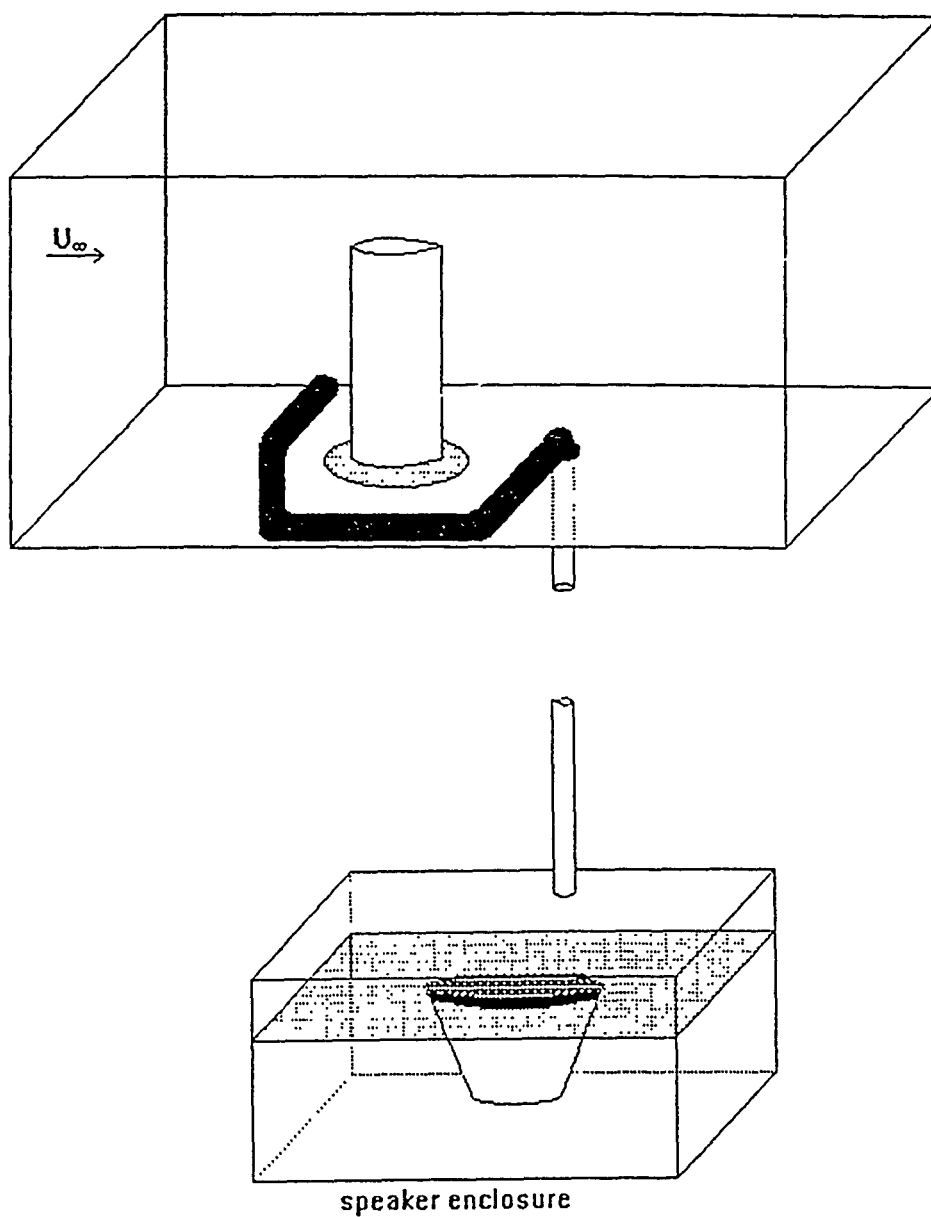
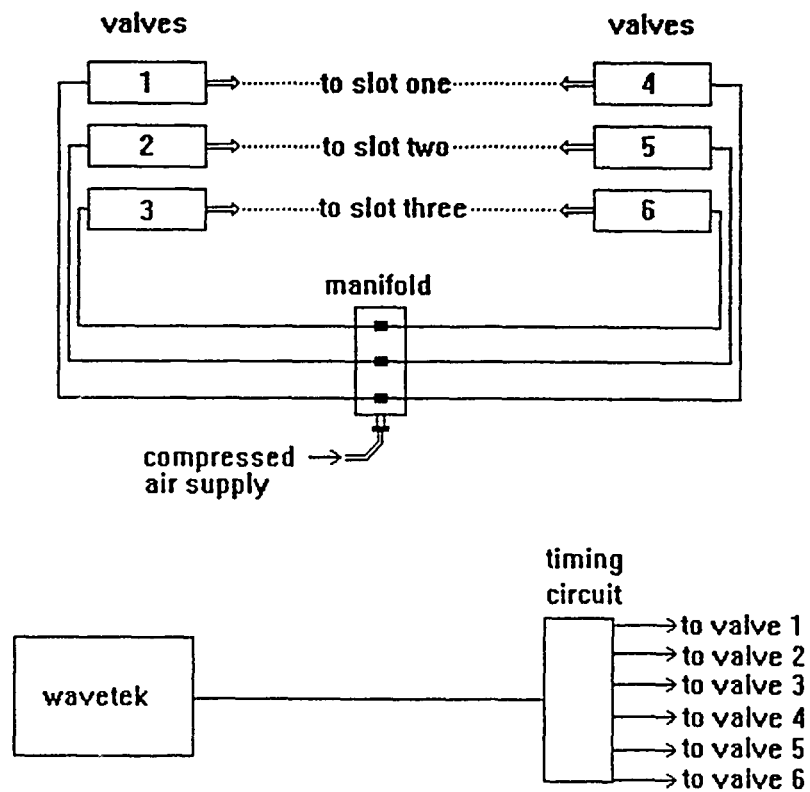
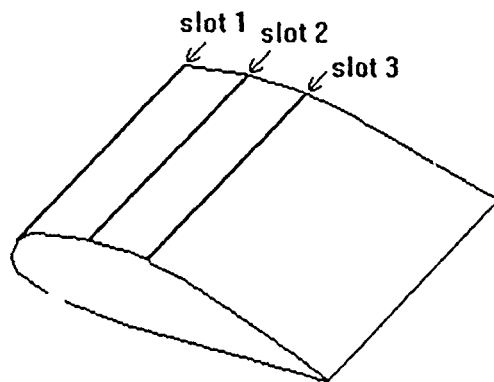


Figure 3. Acoustic Control Configuration.



PULSED-AIR AIRFOIL CONFIGURATION



SLOT DETAIL

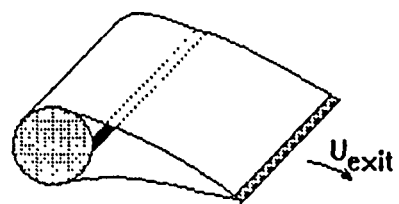


Figure 4. Internal Tangential Pulsed Air Forcing System.

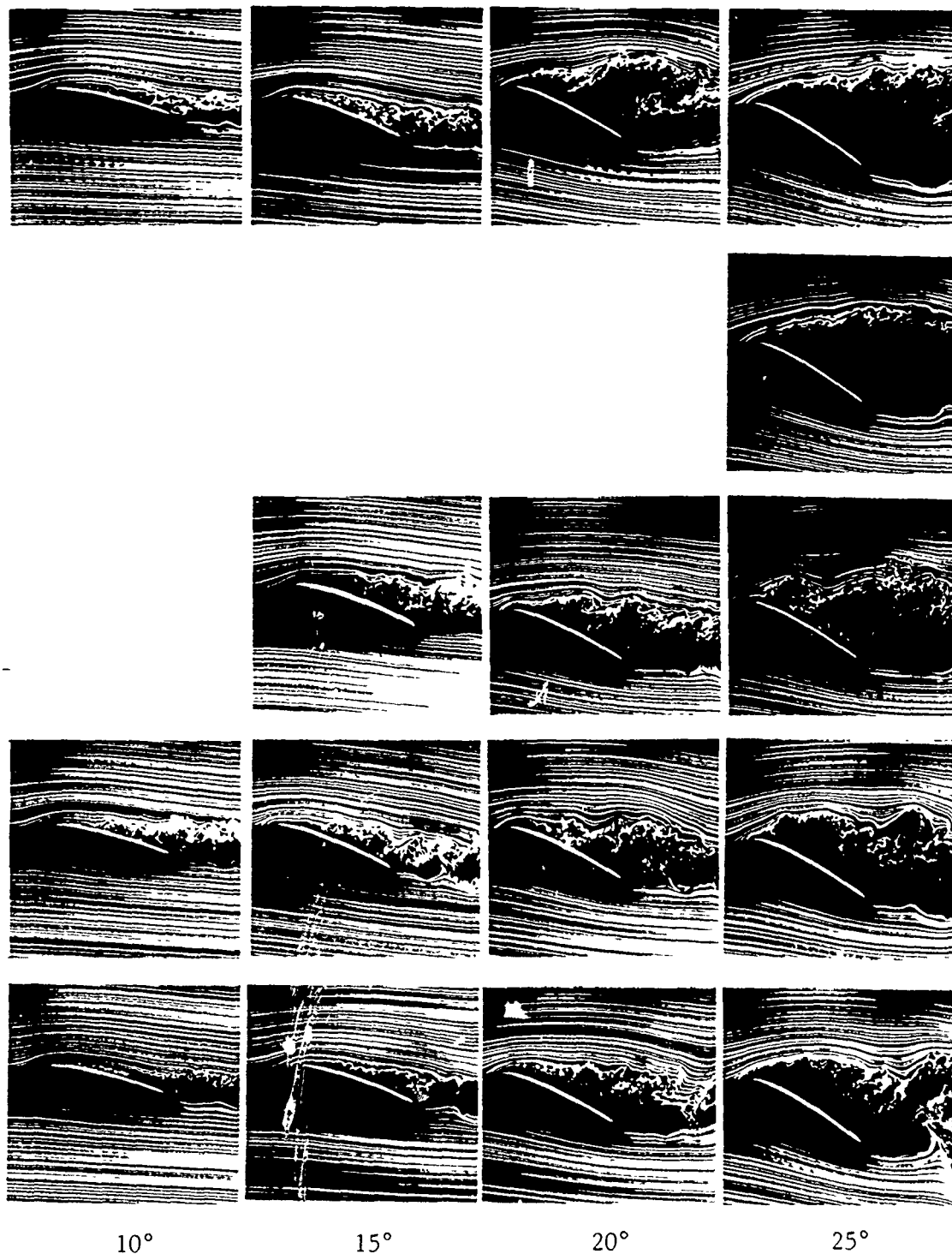
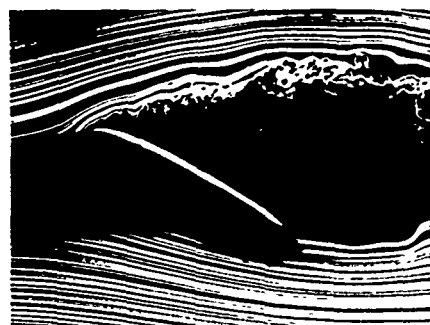
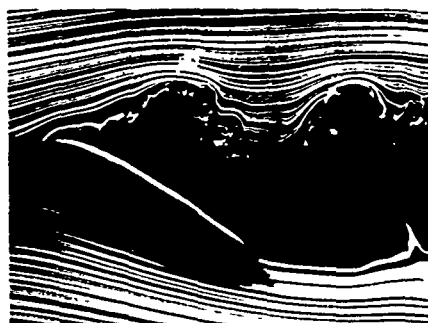


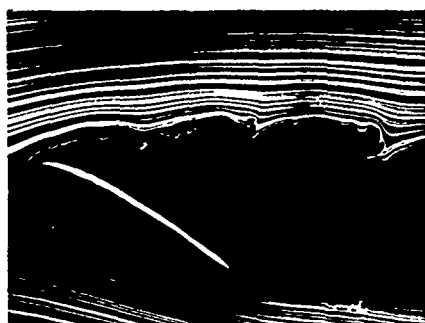
Figure 5. Static NACA-0015 Airfoil under Pulsed Air Control.
 Top row = natural flow, second row = fundamental frequency
 third row = first subharmonic, fourth row = second
 subharmonic, fifth row = third subharmonic.



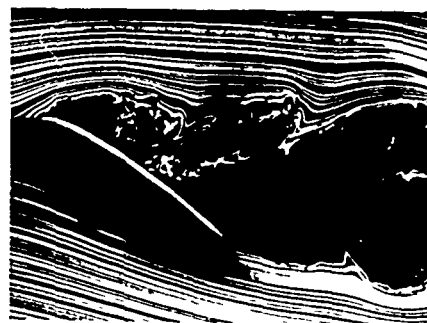
(a) 36 Hz



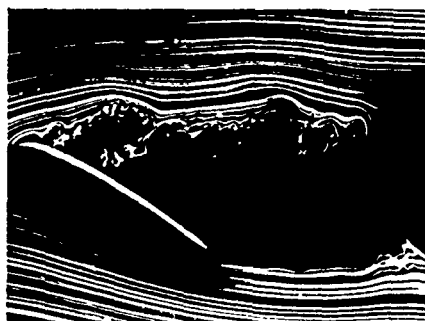
(e) 13 Hz



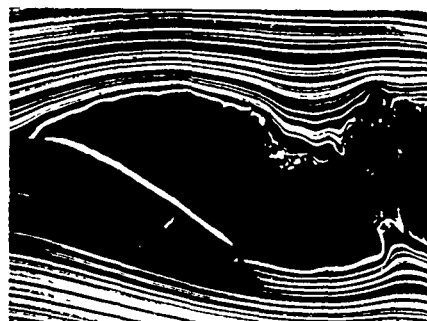
(b) 31 Hz



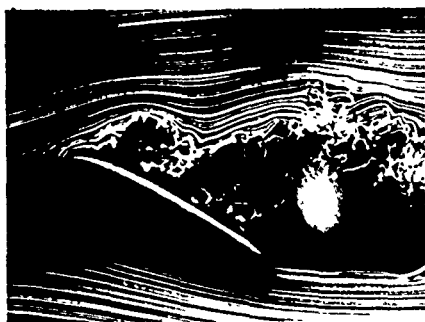
(f) 11 Hz



(c) 22 Hz



(g) 6 Hz



(d) 18 Hz

Figure 6. Static NACA-0015 Airfoil under Pulsed Air Control at 25°.

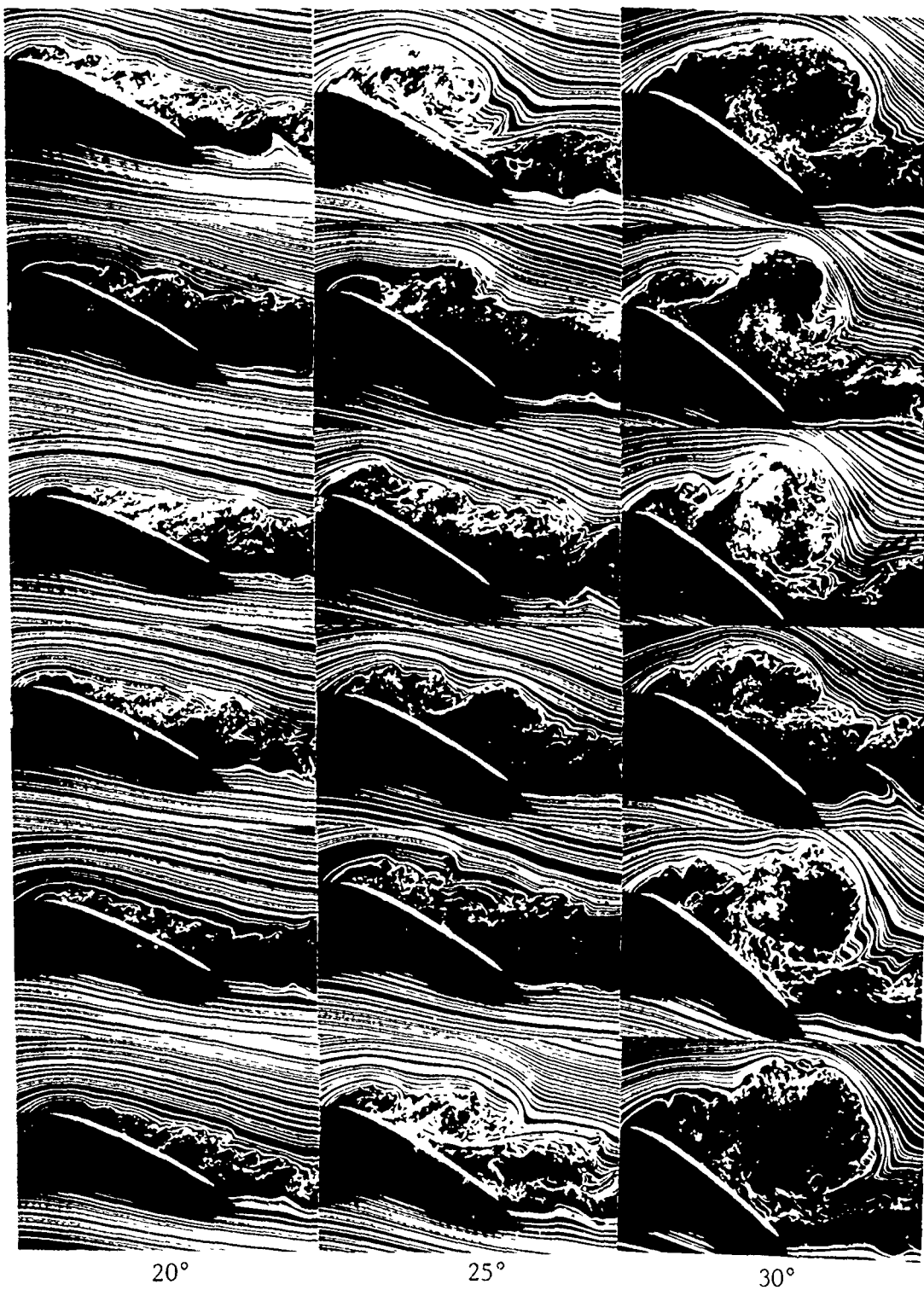


Figure 7. Dynamic NACA-0015 Airfoil under Pulsed Air Control. Top row shows the flow under natural conditions. Each row below that decreases in forcing frequency, in the following order: 22 Hz, 18 Hz, 11 Hz, 9 Hz, 6 Hz.

Relaxation Studies of Molten Salts

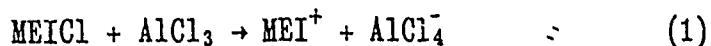
Maureen Parrish

ABSTRACT

The combination of 1-methyl-3-ethylimidazolium chloride, MEICl, with aluminum chloride, AlCl₃, produces a molten salt that has the possibility of being used as a battery electrolyte. In recent years this melt has been used in studying organic solutes.^{1,2} In the present work, ethyl aluminum dichloride, EtAlCl₂, is being added to the melts for the purpose of using the ethyl group as a "environmental NMR probe". The movement of the ethyl group relative to the other melt components will provide useful information and give a better understanding of molten salts as electrochemical and solvent support systems. ¹³C NMR relaxation studies have been done on the seven carbon peaks of these melts. Correlation studies with ²⁷Al NMR are being done to determine the quadrupole coupling constants.

INTRODUCTION

The MEICl and AlCl₃ were made as reported in the literature.³ Once the desired composition of the melt was achieved a predetermined amount of ethyl aluminum dichloride was added.



The melts were made with the mole ratio composition of $n=0.60$ melts of 25, 30, and 35% EtAlCl_2 . In addition a $n=0.50$ melt with only MEICl and EtAlCl_2 was also tested.

$$n_{\text{total}} = 0.60 \quad n_{\text{EtAlCl}_2} = 0.25 \quad n_{\text{AlCl}_3} = 0.35$$

$$\text{gms of AlCl}_3 = \frac{(n_{\text{AlCl}_3}) (W_{\text{total}})}{(n_{\text{AlCl}_3}) - [(MW_{\text{MEICl}}/MW_{\text{AlCl}_3})(n_{\text{total}} - 1)]}$$

$$\text{gms EtAlCl}_2 = \frac{(n_{\text{EtAlCl}_2}) (W_{\text{total}})}{(n_{\text{total}}) - [(MW_{\text{MEICl}}/MW_{\text{EtAlCl}_2})(n_{\text{total}} - 1)]}$$

$$\text{gms MEICl} = (\text{gms AlCl}_3 + \text{gms EtAlCl}_2) - \text{gms of melt}$$

Table I: Moles needed to make 20 grams of melt.

| % EtAlCl_2 | 50% | 25% | 30% | 35% |
|---------------------|-------|-------|-------|-------|
| MEICl | 0.073 | 0.058 | 0.058 | 0.058 |
| AlCl_3 | ---- | 0.050 | 0.043 | 0.036 |
| EtAlCl_2 | 0.073 | 0.036 | 0.043 | 0.050 |

The formation of the melts was done in a helium filled dry box to prevent the reaction of EtAlCl_2 with air. The combination of MEICl and AlCl_3 was not accomplished at any specific time but the addition of EtAlCl_2 was done the day before the testing. The purpose of the timed addition of EtAlCl_2 was to reduce the chance of losing EtAlCl_2 . EtAlCl_2 was added by pipette; the amount was determined by weight, not volume. The initial addition of EtAlCl_2 was done at a very slow rate (0.5

grams/five minutes). This was done to reduce heat that evolves from the reaction of EtAlCl_2 with H_2O . After the white MEICl was no longer visible and a low amount of heat was evolving (determined by touch of gloved hand), a larger amount of EtAlCl_2 could be added at one time (1.0 gram). The samples were loaded into a 5 mm NMR tubes, capped and sealed by a glass blower.

NMR measurements were recorded on a Bruker AC-300 NMR (75.472 MHz) spectrometer. The temperature was calibrated against ethylene glycol or methanol and is accurate to within 0.5°C . Pulse width (^{13}C) was $5.1 \mu\text{s}$, the number of scans was 16, SI was 16K and the delay time was varied from 10 to 37 sec.

NOE (η) measurements were made using the gated decoupler method.⁴ In this method the fully decoupled spectrum is compared with one in which decoupling is present only during the acquisition time. The error in the NOE measurements is approximately 5-10 %.

DISCUSSION

Considerable information about the dynamics and structure of this melt can be obtained through the use of ^{13}C NMR relaxation studies.⁵ The equation used for ^{13}C dipolar relaxation in which the ^{13}C nucleus is relaxed by ^1H is given by:⁶

$$R_1^{\text{dd}} = N_h (h \gamma_c \gamma_h)^2 r_{\text{ch}}^{-6} \tau_{\text{eff}} \quad (3)$$

The dipolar relaxation rate is $R_1^{\text{dd}} = (1/T_1^{\text{dd}})$, N_h is the number of hydrogens attached directly to the carbon atom, γ_c and γ_h are

gyromagnetic ratios and $r_{\text{ch}} = 1.09 \times 10^{-8} \text{ cm}$. The effective correlation time, τ_{eff} , varies exponentially with temperature. Equation (3) is applicable for small molecules including the chloroaluminate/ethyl aluminum dichloride melts. The activation energies (E_a) are indicative of a molecule's ability to move through the medium. E_a is determined from R_1^{dd} as a function of temperature and the results can support the existence of interactions between an ion (MEI^+) and its nearest neighbors (Cl^- , AlCl_4^- , Al_2Cl_7^- , or EtAlCl_3^-).

Isotropic molecular tumbling can be monitored by the use of a NMR "probe" and with a nucleus that relaxes primarily through a quadrupolar mechanism and has a spin greater than $1/2$. Quadrupole interaction results from a distortion of tetrahedral or octahedral symmetry. Nuclei, such as ^{27}Al , is under the influence of an electric field gradient which produces the quadrupole interaction. The quadrupolar relaxation rate in the extreme narrowing region is given by:^{6,7}

$$R_1 = \frac{1}{T_1} = \frac{3\pi^2}{10} \frac{(2I+3)}{I^2(2I-1)} \left[1 + \frac{z^2}{3}\right] \left[\frac{e^2 Q q}{h}\right]^2 \tau_c \quad (4)$$

where $I = 5/2$ for ^{27}Al , eQ is the nuclear electric quadrupole moment, eq is the maximum component of the electric field gradient tensor, and z is the asymmetry parameter of the electric field gradient tensor. (See tables I to V) E_a is equal to the slope times 1.987 (see Fig. 2 to 13). The quadrupole coupling constant (QCC) is the collection of terms given by Equation (5)

$$\text{QCC} = \frac{e^2 Q q}{h} \quad (5)$$

There are a few reports in the literature of QCC for ^{27}Al complexes or compounds in the liquid state. The distortion from octahedral symmetry has produced a range of QCC values such as:

| | |
|-------------------------------|----------|
| alumichrome ⁸ | 32.9 MHz |
| triethylaluminum ⁹ | 23.3 MHz |
| aluminum acetyl ¹⁰ | 0.49 MHz |

These results were obtained by either using the solid state QCC,⁸ assuming an effective correlation time,⁹ or by using a dual spin probe method.¹⁰ The dual spin probe (DSP) method requires that the molecular tumbling is isotropic and that τ_{eff} in eqn. (3) is identical to τ_c in eqn (4). If this is true, then QCC is determined from ^{13}C R_1^{dd} 's and from ^{27}Al R_1 's, assuming that $z = 0$. This last assumption may introduce a small error into the determination of QCC, however, the known tetrahedral symmetry of AlCl_4^- provides good support for the assumption.

It will be assumed that the $\text{MEI}(\text{AlCl}_4)_n^{-(n-1)}$ complexes¹¹ provide a basis for the use of the DSP method. If this assumption is valid, the combination of ^{13}C relaxation data from the nucleus of MEI^+ can be used to determine the QCC for the associated AlCl_4^- species. This is accomplished by plotting the ^{13}C dipolar relaxation rates at each temperature versus the Al relaxation rates over the temperature range 0-70°C. If $\tau_{\text{eff}} = \tau_c$, then from (3) and (4),

$$R_1^{\text{dd}} / N_h (h \gamma_c \gamma_h)^2 \tau_{\text{ch}}^{-6} = R_1 (^{27}\text{Al}) / a \chi^2 \quad (6)$$

$$\text{where } a = \frac{3\pi^2}{10} \frac{(2I+3)}{I^2(2I-1)} \left[1 + \frac{z^2}{3} \right] \quad (7)$$

Consequently, a plot of R_1^{dd} versus R_1 (^{27}Al), will have a zero intercept and a slope, S , equal to (8):

$$S = N_h (h \gamma_c \gamma_h)^2 r_{ch}^{-6} / a \chi^2 \quad (8)$$

or the QCC (χ) equals (9):

$$\text{QCC} = \chi = (h \gamma_c \gamma_h) (N_h / a S)^{1/2} r_{ch}^{-3} \quad (9)$$

Results:

The following tables contain the results of the T_1 measurements for the various melts. Time constraints prevented me from completing the NOE measurements for all of the various melts. However, these measurements are being completed on the 300 MHz NMR Spectrometer located at Wichita State University. Dr. Carper's research group will finish these measurements and calculate the R_1^{dd} for each melt over the temperature range previously specified ($25^\circ - 70^\circ \text{C}$).

Figure 1:

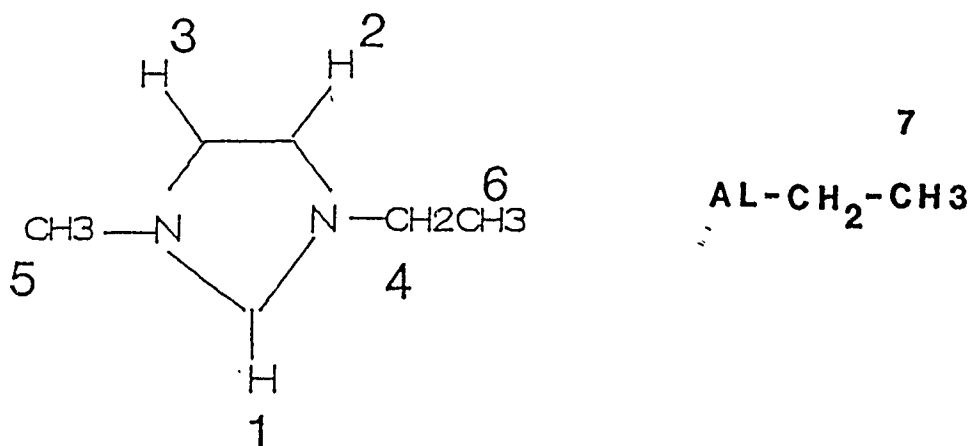


Table II: R_1 's for a 50% MEICl/EtAlCl₂ melt.

| Temp | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|------|------|-------|------|------|------|------|
| 70 | 0.28 | 0.32 | 0.27 | 0.22 | 0.23 | 0.11 |
| 65 | 0.33 | 0.35 | 0.31 | 0.24 | 0.26 | 0.11 |
| 60 | 0.35 | 0.38 | 0.34 | 0.25 | 0.28 | 0.12 |
| 55 | 0.39 | 0.43 | 0.38 | 0.27 | 0.30 | 0.12 |
| 50 | 0.43 | 0.46 | 0.40 | 0.29 | 0.32 | 0.13 |
| 45 | 0.47 | 0.52 | 0.45 | 0.31 | 0.35 | 0.14 |
| 40 | 0.53 | 0.59 | 0.51 | 0.34 | 0.38 | 0.15 |
| 35 | 0.62 | 0.68 | 0.57 | 0.37 | 0.43 | 0.16 |
| 30 | 0.71 | 0.77 | 0.63 | 0.40 | 0.47 | 0.17 |
| 25 | 0.77 | 0.86 | 0.70 | 0.43 | 0.50 | 0.18 |

Table III: $-\ln R_1$'s for a n=0.50 melt MEICl/EtAlCl₂

| Temp. | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|-------|------|-------|------|------|------|------|
| 70 | 1.27 | 1.12 | 1.32 | 1.52 | 1.47 | 2.23 |
| 65 | 1.12 | 1.04 | 1.17 | 1.42 | 1.35 | 2.20 |
| 60 | 1.05 | 0.96 | 1.07 | 1.37 | 1.28 | 2.15 |
| 55 | 0.94 | 0.85 | 0.98 | 1.29 | 1.18 | 2.11 |
| 50 | 0.85 | 0.77 | 0.91 | 1.25 | 1.14 | 2.04 |
| 45 | 0.75 | 0.64 | 0.80 | 1.18 | 1.06 | 1.97 |
| 40 | 0.63 | 0.52 | 0.68 | 1.08 | 0.96 | 1.90 |
| 35 | 0.47 | 0.39 | 0.57 | 0.99 | 0.84 | 1.81 |
| 30 | 0.34 | 0.26 | 0.47 | 0.91 | 0.75 | 1.74 |
| 25 | 0.27 | 0.15 | 0.35 | 0.83 | 0.68 | 1.71 |

Table IV: R_1 's for a 35% EtAlCl₂ n = 0.60 melt.

| Temp | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|------|------|-------|------|------|------|------|
| 70 | 0.22 | 0.25 | 0.24 | 0.21 | 0.22 | 0.10 |
| 65 | 0.23 | 0.27 | 0.26 | 0.22 | 0.23 | 0.12 |
| 60 | 0.25 | 0.30 | 0.28 | 0.23 | 0.25 | 0.11 |
| 55 | 0.28 | 0.33 | 0.31 | 0.24 | 0.27 | 0.12 |
| 50 | 0.32 | 0.37 | 0.34 | 0.25 | 0.29 | 0.13 |
| 45 | 0.36 | 0.41 | 0.38 | 0.29 | 0.32 | 0.13 |
| 40 | 0.40 | 0.46 | 0.41 | 0.30 | 0.36 | 0.15 |
| 35 | 0.45 | 0.51 | 0.47 | 0.32 | 0.37 | 0.16 |
| 30 | 0.50 | 0.58 | 0.53 | 0.36 | 0.42 | 0.17 |
| 25 | 0.56 | 0.64 | 0.57 | 0.38 | 0.44 | 0.18 |

Table V: $-\ln R_1$'s for a 35% EtAlCl₂ n=0.60 melt

| Temp | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|------|------|-------|------|------|------|------|
| 70 | 1.49 | 1.37 | 1.43 | 1.55 | 1.50 | 2.29 |
| 65 | 1.48 | 1.29 | 1.33 | 1.50 | 1.49 | 2.23 |
| 60 | 1.38 | 1.21 | 1.26 | 1.47 | 1.39 | 2.16 |
| 55 | 1.26 | 1.10 | 1.16 | 1.41 | 1.31 | 2.11 |
| 50 | 1.15 | 0.99 | 1.07 | 1.37 | 1.25 | 2.03 |
| 45 | 1.02 | 0.88 | 0.96 | 1.28 | 1.15 | 2.00 |
| 40 | 0.92 | 0.78 | 0.88 | 1.21 | 1.03 | 1.89 |
| 35 | 0.80 | 0.67 | 0.76 | 1.15 | 0.99 | 1.83 |
| 30 | 0.69 | 0.55 | 0.63 | 1.03 | 0.86 | 1.75 |
| 25 | 0.58 | 0.45 | 0.57 | 0.97 | 0.81 | 1.70 |

Table VI: R_1 's for a 25% EtAlCl₂ n = 0.60 melt.

| Temp | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|------|------|-------|------|------|------|------|
| 70 | 0.21 | 0.25 | 0.24 | 0.21 | 0.23 | 0.09 |
| 65 | 0.23 | 0.27 | 0.26 | 0.21 | 0.24 | 0.11 |
| 60 | 0.26 | 0.30 | 0.28 | 0.22 | 0.26 | 0.11 |
| 55 | 0.29 | 0.34 | 0.31 | 0.24 | 0.28 | 0.12 |
| 50 | 0.32 | 0.37 | 0.34 | 0.25 | 0.29 | 0.12 |
| 45 | 0.35 | 0.41 | 0.38 | 0.27 | 0.32 | 0.13 |
| 40 | 0.40 | 0.46 | 0.41 | 0.30 | 0.35 | 0.14 |
| 35 | 0.44 | 0.51 | 0.45 | 0.32 | 0.38 | 0.15 |
| 30 | 0.50 | 0.58 | 0.51 | 0.33 | 0.41 | 0.16 |
| 25 | 0.55 | 0.63 | 0.57 | 0.36 | 0.44 | 0.18 |

Table VII: $-\ln R_1$'s for a 25% EtAlCl₂ n=0.60 melt

| Temp | pk1 | pk2/3 | pk4 | pk5 | pk6 | pk7 |
|------|------|-------|------|------|------|------|
| 70 | 1.55 | 1.38 | 1.43 | 1.60 | 1.48 | 2.34 |
| 65 | 1.45 | 1.29 | 1.35 | 1.55 | 1.47 | 2.22 |
| 60 | 1.36 | 1.19 | 1.27 | 1.49 | 1.35 | 2.21 |
| 55 | 1.25 | 1.09 | 1.18 | 1.43 | 1.26 | 2.15 |
| 50 | 1.15 | 0.99 | 1.09 | 1.36 | 1.21 | 2.11 |
| 45 | 1.06 | 0.89 | 0.97 | 1.29 | 1.13 | 2.02 |
| 40 | 0.92 | 0.78 | 0.89 | 1.22 | 1.04 | 1.94 |
| 35 | 0.81 | 0.67 | 0.80 | 1.15 | 0.97 | 1.87 |
| 30 | 0.70 | 0.55 | 0.68 | 1.10 | 0.89 | 1.80 |
| 25 | 0.60 | 0.46 | 0.57 | 1.01 | 0.81 | 1.70 |

Table IIV: E_a (Kcal) for various melts.

| % EtAlCl ₂ | pk1 | pk 2 | pk 4 | pk 5 | pk 6 | pk7 |
|--------------------------|------|------|------|------|------|------|
| 50 | 4.96 | 4.51 | 4.21 | 3.04 | 3.50 | 2.48 |
| 35 | 4.37 | 4.25 | 3.97 | 2.70 | 3.34 | 2.72 |
| 30 | 4.19 | 4.15 | 3.74 | 3.00 | 3.20 | 3.12 |
| 25 | 4.41 | 4.21 | 3.87 | 2.76 | 3.04 | 2.90 |

Conclusion:

The main relationship that can be derived from these sets of data are concerned with the E_a over the temperature range 25 - 70°C. Examination of table IV reveals definite trends associated with changes in melt composition. It is apparent that the molecular association between MEICl and EtAlCl₂ is considerably weaker than the association of AlCl₃ with MEICl. This conclusion is reinforced by the decrease in E_a as a function of melt composition (i.e., decrease in AlCl₃ content). Peaks 5 (CH₃-N), 6 (CH₃-CH₂N), and 7 (CH₃-CH₂Al) have the lower activation energy since they have more flexibility and undergo relaxation via additional mechanisms. Peaks 1 (CH₂ between the nitrogens), 2/3 (CH=CH), and 4 (CH₂-N) are fixed in their position and require more energy to be released from the complex.

In summary, the MEICl carbon ring has a maximum E_a due to the breakup of the complex whereas the other carbons are subject to internal rotation and are more loosely bound. The lowest E_a values are consistent with maximum internal rotation.

Bibliography

- 1) Wilkes, J.S.; Levisky, J.A.; Wilson, R.A.; Hussey C.L., *Inorg. Chem* 1982, 21, 1263.
- 2) Dymek C.J., Jr. and G.E. Godec, Viscosity and Conductance of MEIC/ AlCl_3 Melts Containing Nonaqueous Solvents, FJSRL-TR-88-0003, May 1988.
- 3) Parrish, Maureen, The Preparation and Purification of Melt: 1-Methyl-3-Ethylimidazolium Chloride (MeEtImCl) and Aluminum Chloride (AlCl_3), in preparation.
- 4) Neuhaus, D.; Williamson, M. "The Nuclear Overhauser Effect in Structural and Conformational Analysis", VCH Publishers: New York, 1984; p. 56.
- 5) Carper, W.R.; Pflug, J.L.; Elias, A.M.; Wilkes, J.S, *J. Amer. Chem. Soc.* (submitted).
- 6) Abragam, A. "Principles of Nuclear Magnetism", Oxford University Press: Oxford, 1961; Chap. 8.
- 7) Lindmen, B.; Forsen, S.; "NMR Basic Principles and Progress"; Diehl, P.; Pluck, E.; Kosfield, R., Ed.; Vol. 12, Springer-Verlag: NewYork, 1976, p.22.
- 8) Llinas, M.; DeMarco, A.J., *J. Amer. Chem. Soc.* 1980, 102, 2226.
- 9) Vestin, V.; Dowalewoki, J.; Henriksson, U., *Org. Magn. Res.* 1981, 16, 119.
- 10) Dechter, J.J.; Henrikson, U.J., *Magn. Res..* 1982, 48, 503.
- 11) Carper, W.R.; Pflug, J.L.; Wilkes, J.S. J., *Org. Chem.* (submitted).

Fig. 2

.60 MELT (25% EtAlCl₂)

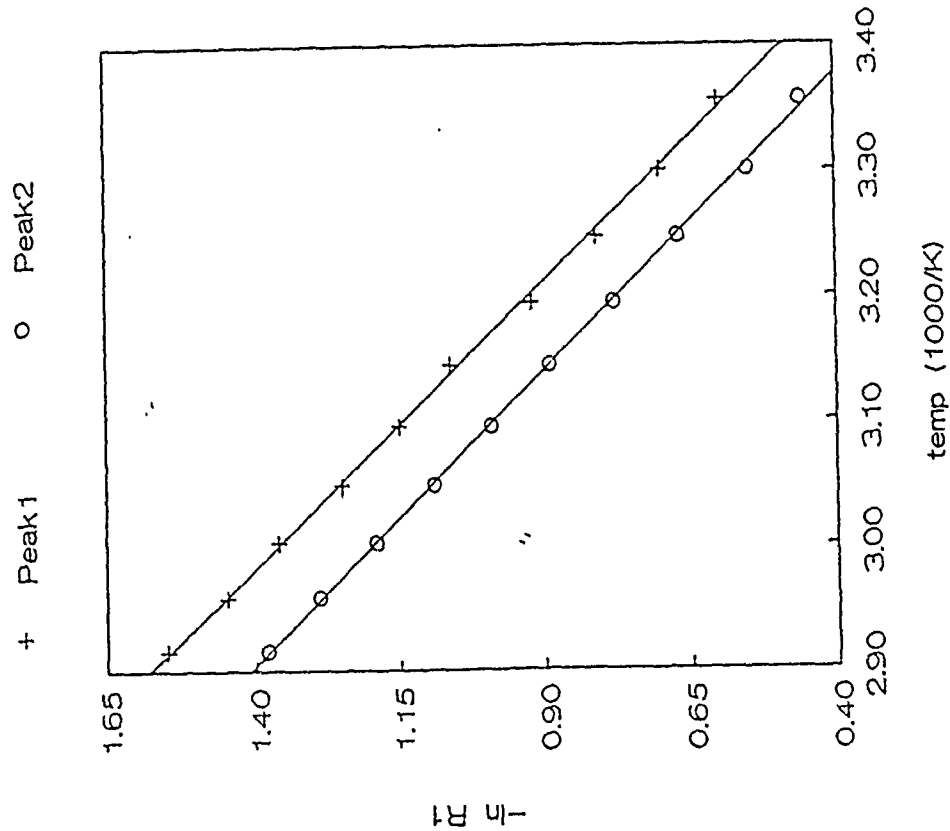


Fig. 3

.60 MELT (25% EtAlCl₂)

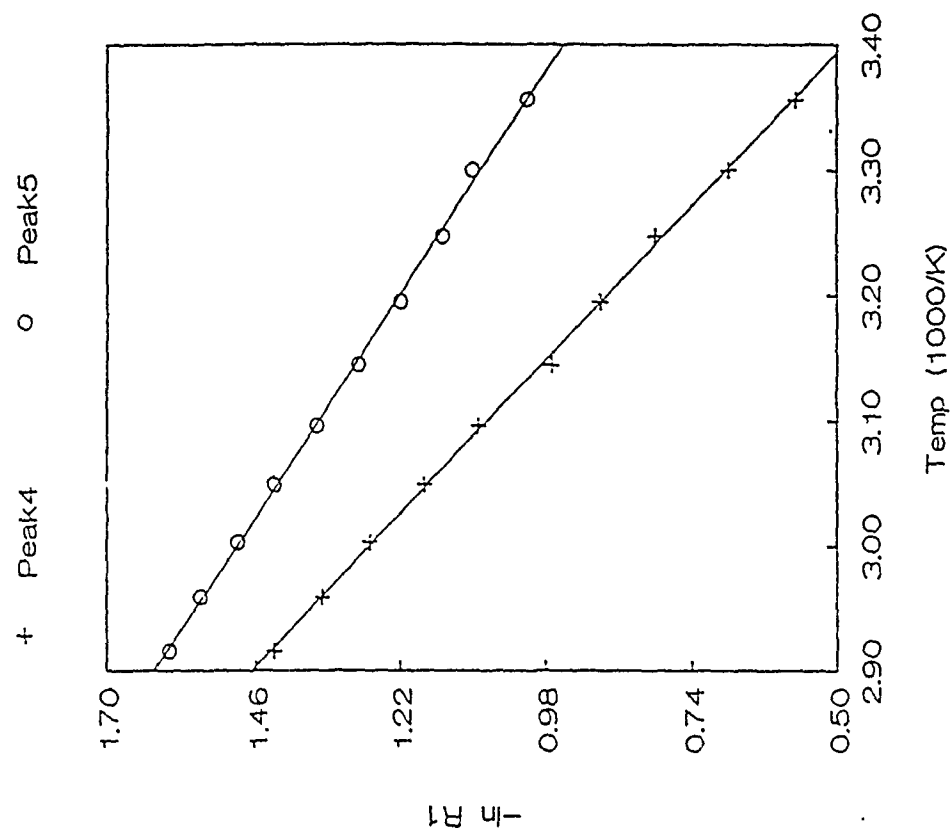


Fig. 4

.60 MELT (25% EtAlCl₂)

+ Peak6 o Peak7

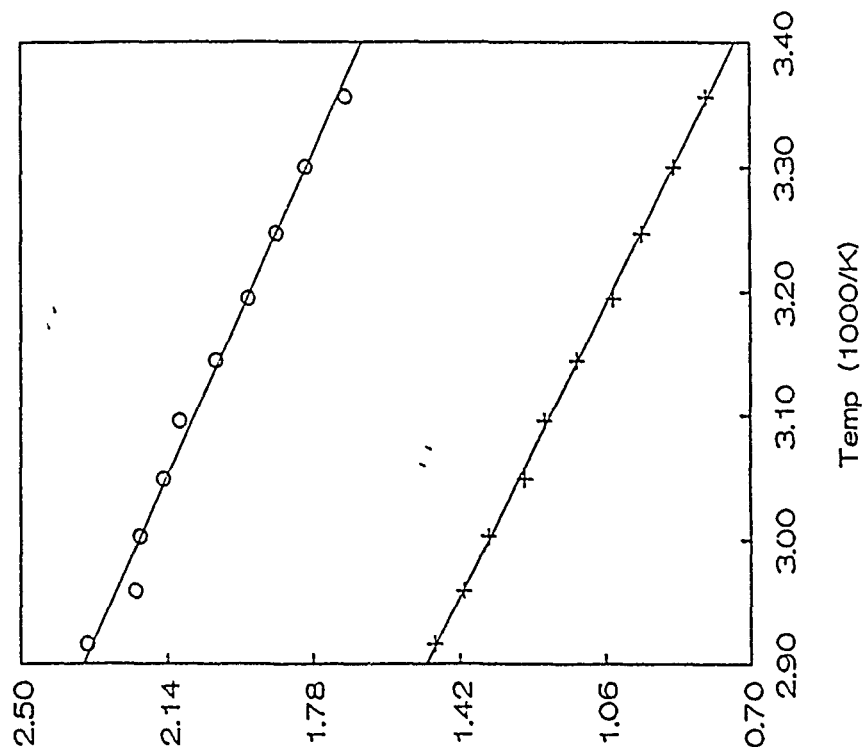


Fig. 5

.60 MELT (30% EtAlCl₂)

+ Peak1 o Peak2

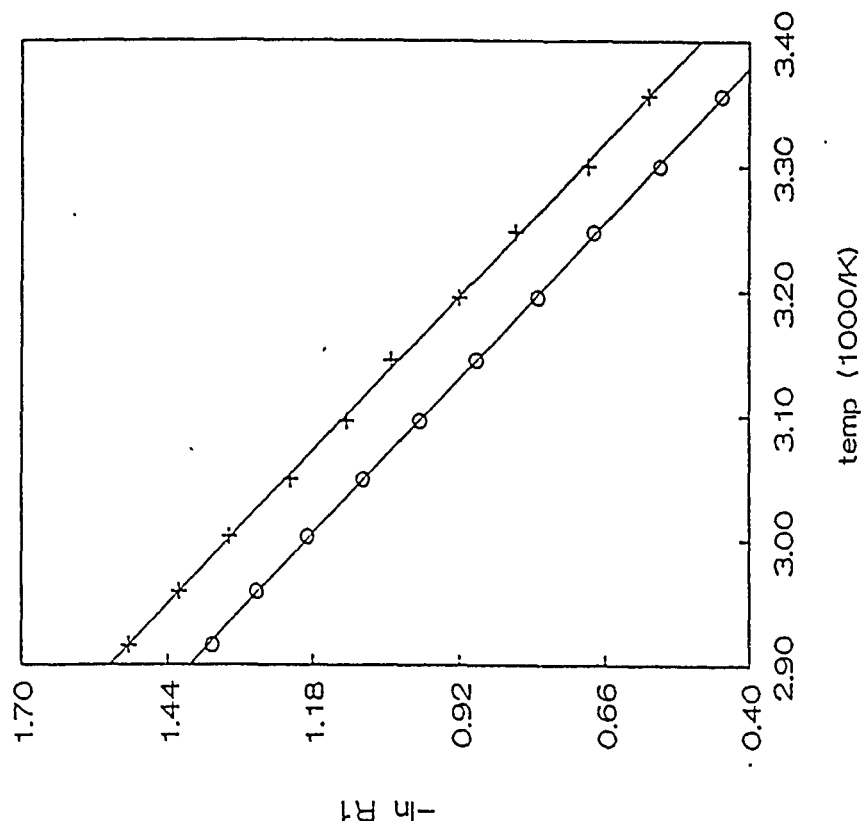


Fig. 6
.60 MELT (30% EtAlCl₂)

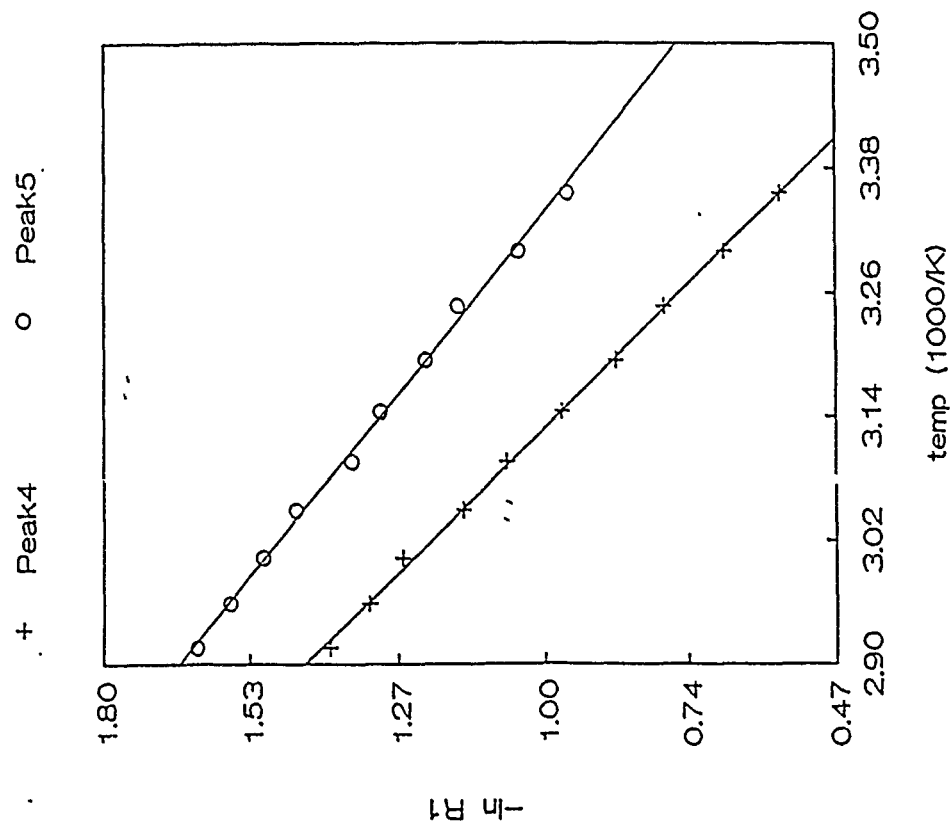


Fig. 7
.60 MELT (30% EtAlCl₂)

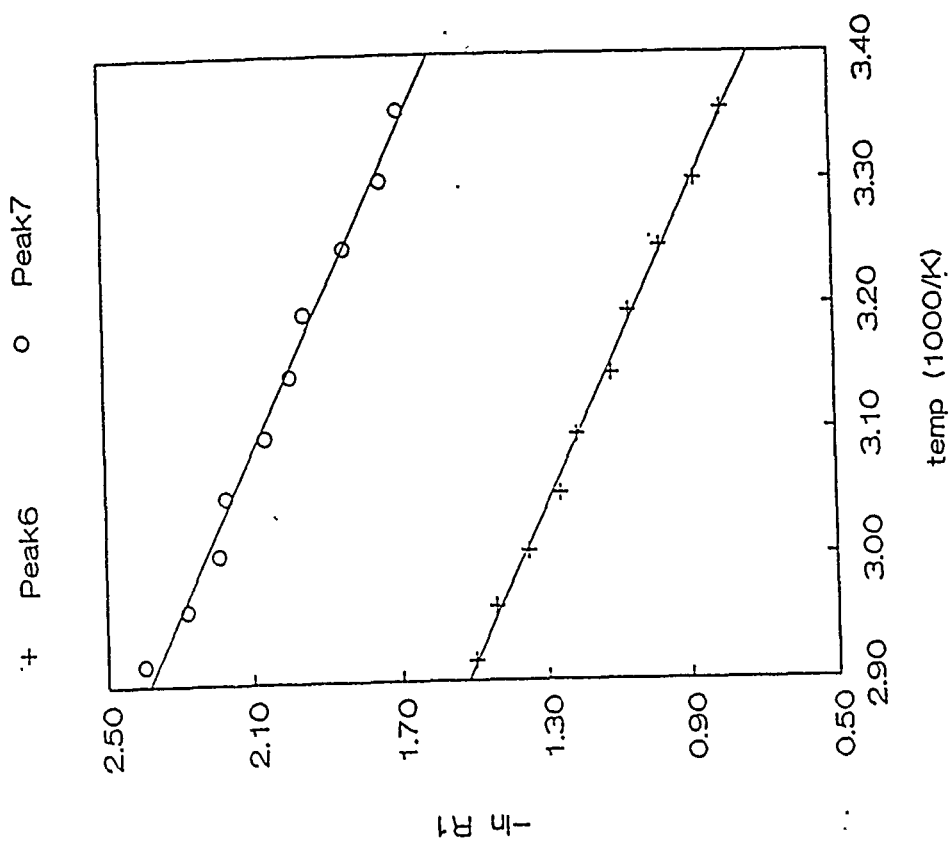


Fig. 8

.60 MELT (35% EtAlCl₂)

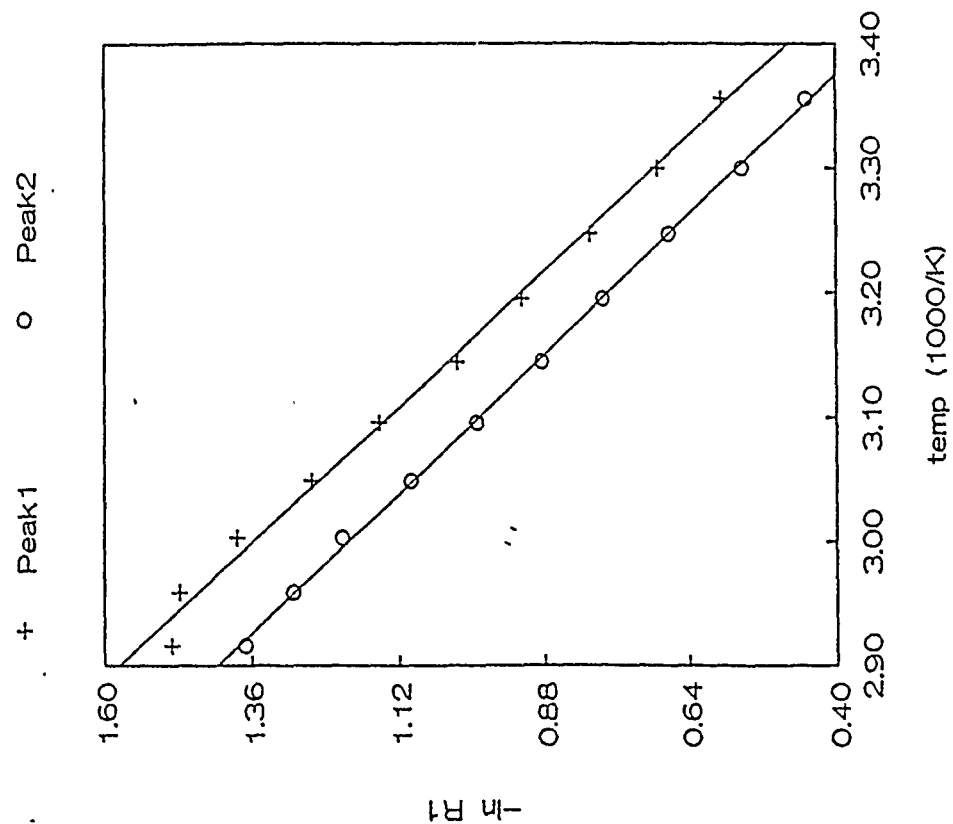


Fig. 9

.60 MELT (35% EtAlCl₂)

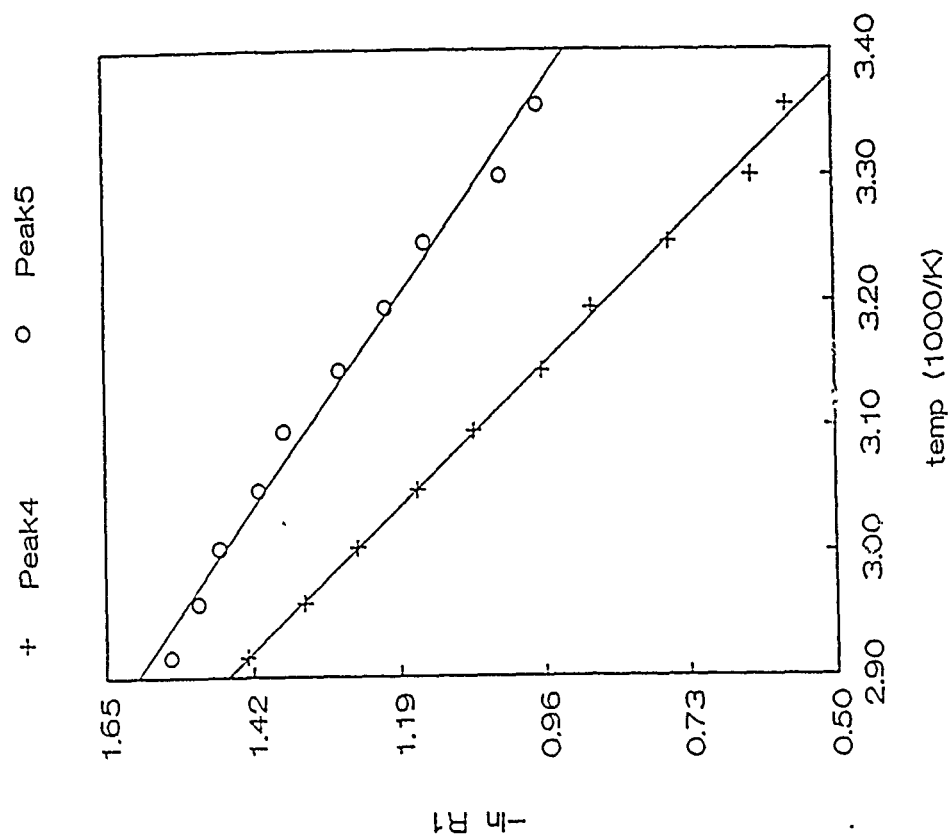


Fig. 10
60 MELT (35% EtAlCl₂)

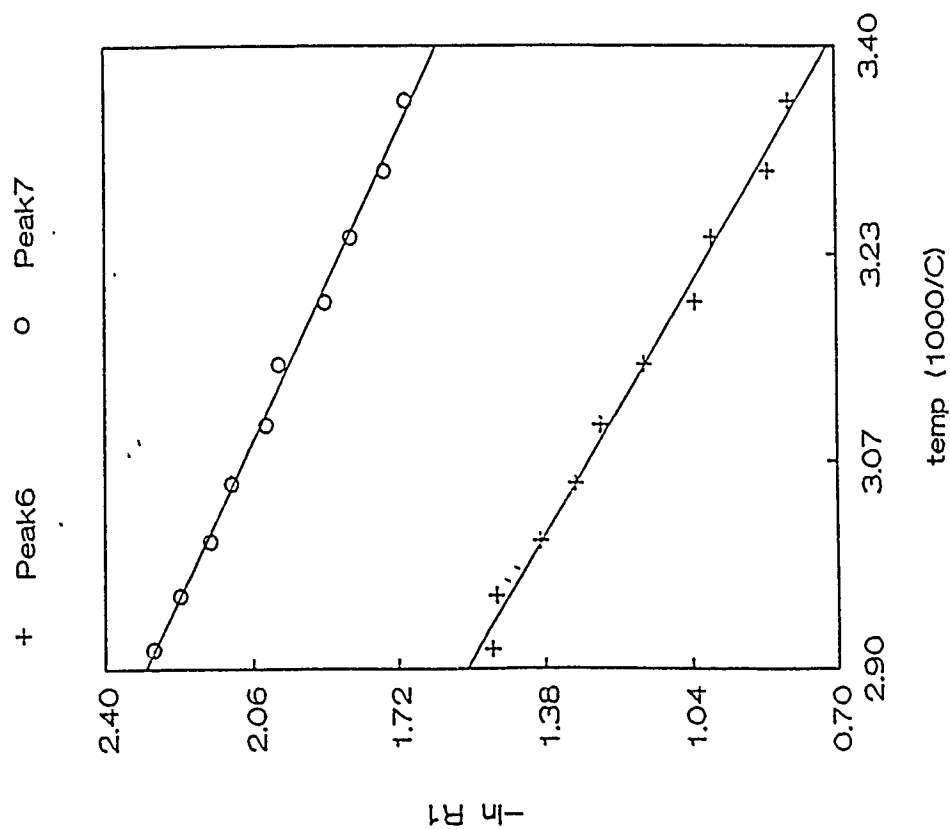


Fig. 11
50 MELT (50% EtAlCl₂)

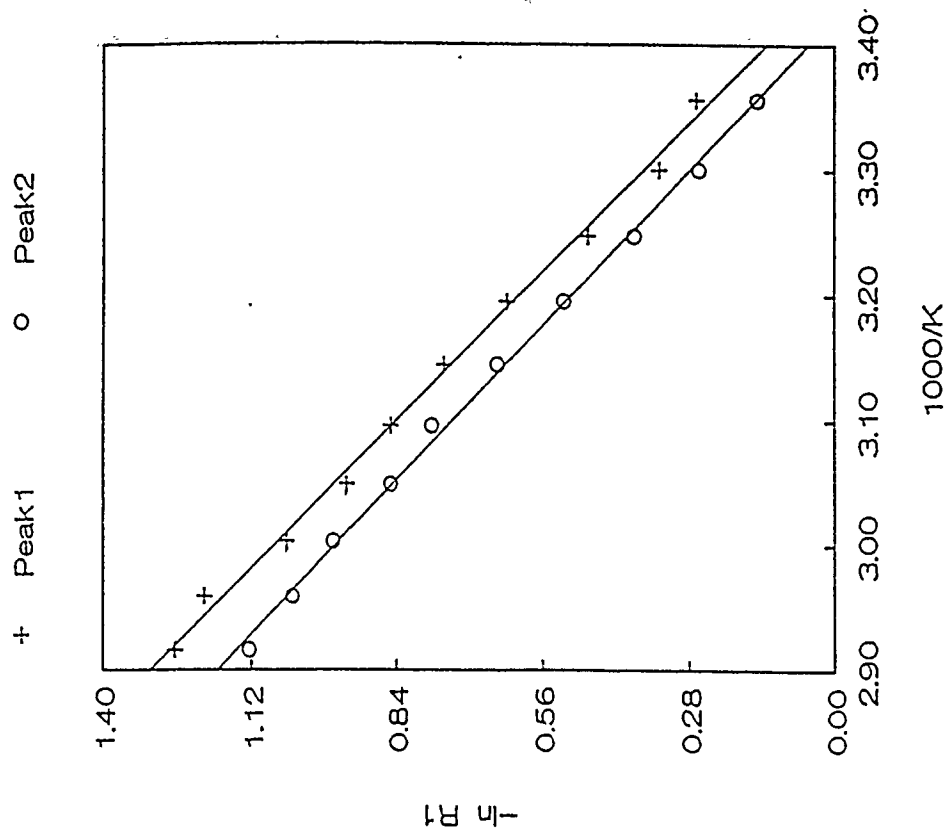


Fig. 2
.5C MELT (50% EtAlCl₂)

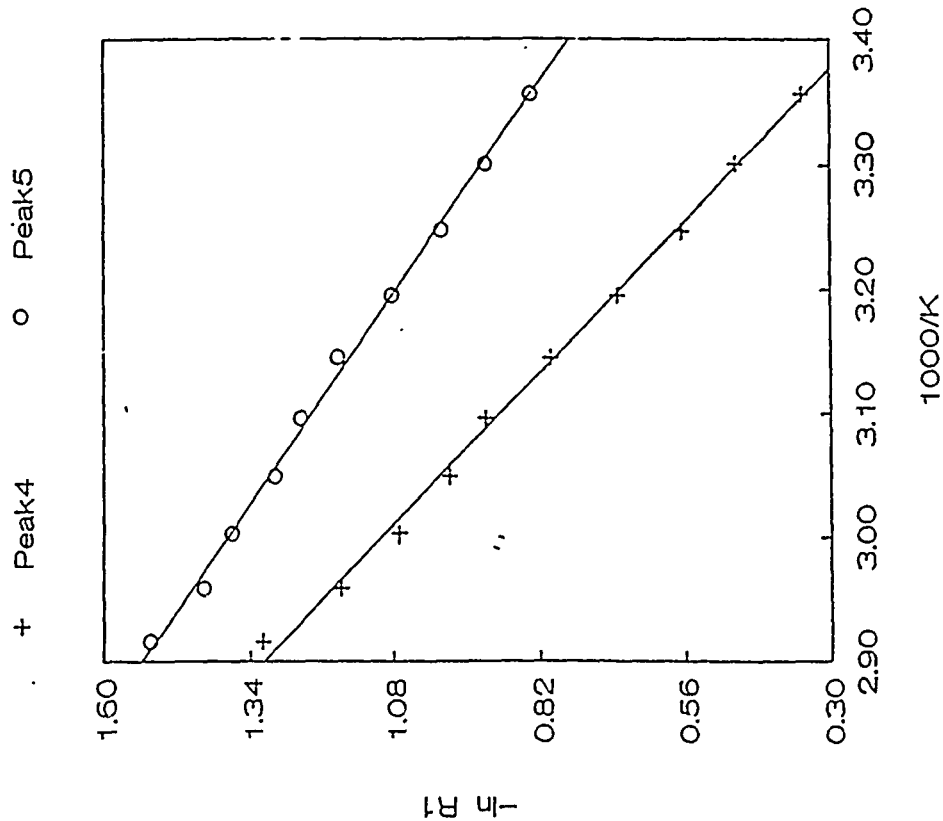
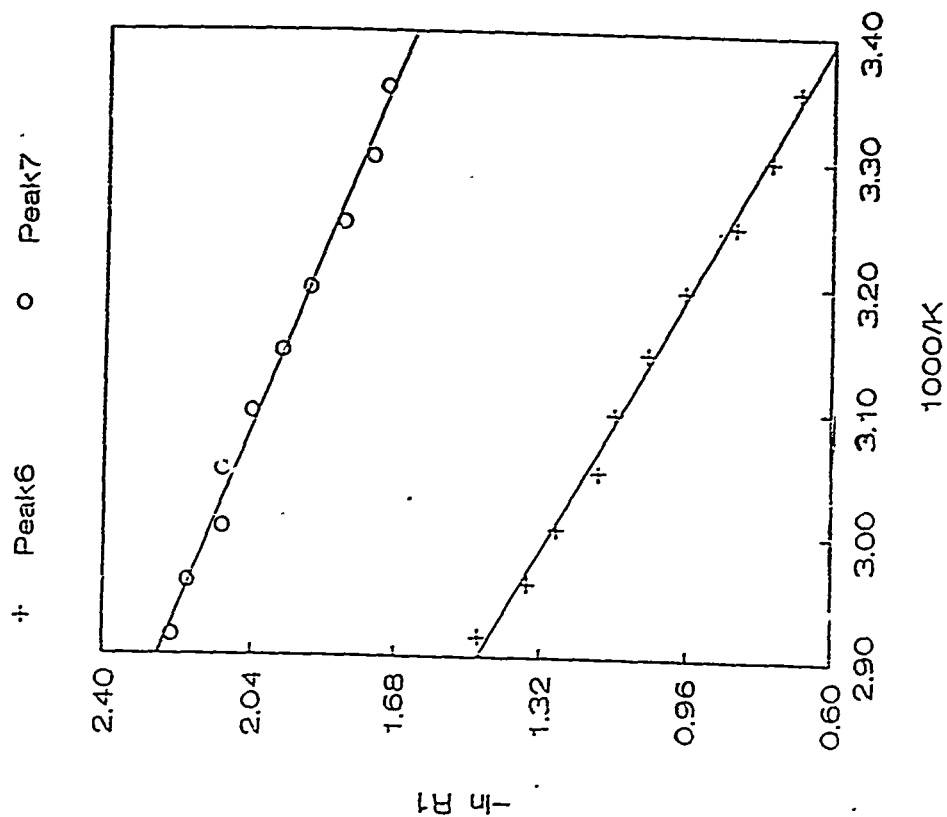


Fig. 13
.50 MELT (50% EtAlCl₂)



DOWNWASH FLOW MECHANISMS ON A PITCHING CANARD/WING CONFIGURATION

Graduate Research Assistant

Dennis J. Strickland Jr.

Abstract

In recent years, much attention has been given to enhancing the maneuverability of fighter aircraft. The understanding of the unsteady flow fields in which these fighter aircraft operate is the key to this enhancement. The primary concern of the present investigation was to correlate flow visualization data with existing unsteady quantitative data for a canard/wing configuration. Using a smoke-wire flow visualization technique, the flow mechanisms which cause downwash were identified. From this analysis, it was determined that a forward canard at positive angles of attack can produce an upwash on an aft located wing. This lift enhancing property was due to the location of shed vortices with respect to the aft wing and the circulation lag associated with pitching airfoils. The flow mechanisms associated with the circulation lag were identified, but the results were inconclusive.

Introduction

Recently, much attention has been given to enhancing the low speed agility of super maneuverable aircraft. The need for fighter aircraft to fly at very high angles of attack with increased maneuverability is fast becoming a realization. Through the proper use of forced unsteady aerodynamic flows, low-speed handling characteristics of fighter aircraft may be dramatically improved; however, a complete understanding of this time dependent flow has yet to be made.

In the past, it has been proven that oscillating wings may produce normal forces and pitching moments at stall that are considerably higher than the static limit.¹ These same unsteady aerodynamic loads are seen to develop on fighter aircraft maneuvering at very

high angles of attack. Tail surface failures have been known to occur in some fighter aircraft operating in these regimes. For this reason, it is important to understand how these loads develop.

It is uncertain as to how the shed vorticity from a forward lifting surface affects the loads on an aft wing. That is, is it possible for downwash from the forward lifting surface to actually become upwash on the aft surface? If so, what are the mechanisms involved in producing upwash? Since few aerodynamicists have investigated downwash from forward lifting surfaces in unsteady flows, it stands to reason that very little knowledge exists in this area.

One aspect of the unsteady flow field that is unique only to these flows, is the production of leading and trailing edge vortices from a lifting surface. These vortices are unlike any vorticity produced in steady state flows. As their name suggests, the leading and trailing edge vortices are shed from the wing leading and trailing edge, respectively. One characteristic of these vortices, is the fact that both vortices are shed parallel to the leading and trailing edge of the lifting surface respectively. How this and other properties of the vortices affect downwash has yet to be understood; however, a complete understanding is imperative in eventually applying the technology of unsteady aerodynamics to supermaneuverable aircraft.

Investigation of the fluid mechanisms associated with dynamic stall has been studied by several researchers; however, there appears to be conflicting opinions as to the fluid mechanics of the shed vortices. From Carr's description in Ref. 1, when the vortex finally moves past the trailing edge, a sudden loss in lift occurs, represented by a spike in the C_L alpha curve. This sudden loss in lift is not coincident with the sudden change in pitching moment about the quarter chord. The vortex movement over the airfoil surface is primarily the cause of these changes in lift and moment.

A second description of the vortex movement is offered in Ref. 3. Flow visualization (smoke-wire) accompanied by pressure measurements on the airfoil surface were

techniques used to derive the conclusions in this report. Albertson observed that at the beginning of the stall sequence, the leading edge vortex first appears on the upper surface of the airfoil. This vortex grows in an elongated manner until it passes over the 25 to 30 percent chord point. The vortex then begins to grow in a direction normal to the airfoil surface while becoming more circular in the process. Prior to this circular growth, lift reaches a maximum and continues to decrease as the vortex moves away from the airfoil, but not necessarily at the airfoil trailing edge as Carr suggests.

Because of the important role that vortex convection and interaction has in unsteady flows, there is a need to resolve conflicting results. This will only come about after many quantitative and qualitative experimental tests. Hopefully, the present tests will give insight into the capabilities that these flow fields have for improving flight characteristics of fighter aircraft.

Using a smoke-wire flow visualization technique, a qualitative analysis of the flow structure surrounding a canard/wing configuration was investigated. This provided a means for observing how these aforementioned vortices interact with one another. The flow visualization data was recorded using a Kodak Ektapro 1000 Imager. This high speed camera has the capability of digitizing the unsteady flow field and displaying the image on a monitor where a frame by frame analysis of the flow field could then be accomplished. With this analysis, the position and size of the leading and trailing edge vortices should be seen throughout the pitching cycle as well as a qualitative answer to the upwash or downwash question.

Besides downwash, generated by shed vorticity, the circulation lag associated with oscillating wings is an important aspect of the unsteady flow field. In this report, downwash is defined as a negative normal force exerted on the aft wing from the shed vorticity of the forward wing. This definition can also be applied to steady state flows; however, downwash does not exhibit the same characteristics in the unsteady regime. It was shown in Ref. 2 that the unsteady downwash curve is shifted in angle of attack from

the steady state curve. This shift is due to the circulation lag associated with the unsteady flow field. According to Burkharter in Ref. 2, there seems to be a correlation between the shift in the downwash curve and the circulation lag on the aft wing. "It is clear, in an oscillating two wing configuration ... the shift in the downwash curve is more or less synchronized with the circulation lag associated with the aft wing."² The circulation lag associated with the aft wing seems to be independent of the shed vorticity of the forward wing. Depending on the strength and amount of interaction of the leading and trailing edge vortices of the forward wing, the lifting properties of the aft wing can be greatly enhanced. One of the goals of this experiment is to explain the shift of the downwash curve, or circulation lag, with vortex convection and interaction, and to be able to see the vortex interactions which cause downwash and circulation lag.

Apparatus and Procedure

Flow visualization techniques are often quite frustrating in that it is difficult to obtain satisfying results, especially in a dynamic analysis. The primary components of the flow visualization system are relatively easy to control and set up; however, in an unsteady regime, proper lighting becomes a problem due to oscillating shadows and difficulty in collimating the light source on the smoke stream. Also, a high speed camera is imperative in order to obtain a proper time history of the unsteady phenomenon; however, resolution of the photographs is sacrificed for the high speed.

For this experiment, two high intensity light sources were arranged so that one source was placed on each side of the clear Plexiglass test section. Because of this setup, the light had a tendency to 'splash' about the test section causing a glare on the photographic image. In order to alleviate this problem, both sides of the test section were covered with poster paper except for a two inch strip oriented parallel with the flow axis (See Fig. 1). This allowed a narrow beam of light to enter the test section and illuminate the thin smoke

stream; consequently, this solved much of the glare problem. To minimize the glare further, the beveled splitter plate and the tips of each wing were painted black. This provided a background which contrasted with the illuminated smoke stream thus providing a more visible flow on the photographic image. One problem that did occur from the glare was the inability to obtain clear images near the root of the wing. The clearest images of the flow field were obtained from the mid-span out to the tip of the wing.

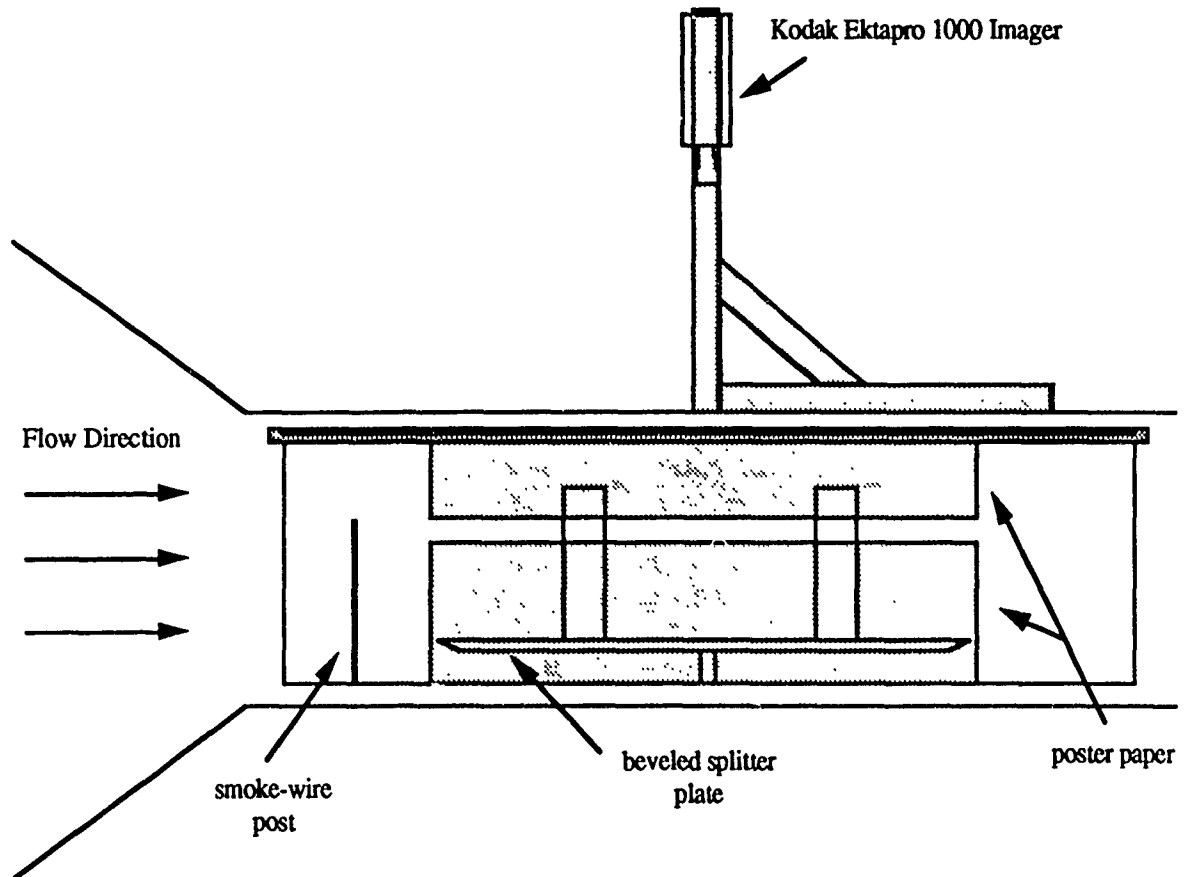


Figure 1: Flow Visualization Set Up

In order to generate the smoke stream, a 0.005 inch diameter tungsten wire was coated with oil then heated via an electric current. A liberal amount of motor oil must be applied to the wire ensuring that small beads form continuously over the wire. If the wire is unevenly coated, there will be irregularities in the smoke stream.

The most important piece of equipment in flow visualization is the high speed camera. Changes occur within a flow field at such a fast rate that many frames per second must be photographed in order to ensure an effective analysis. A Kodak Ektapro 1000 Imager was utilized for this purpose, which had an image rate capability of 30 to 1000 frames per second. Since the test section velocity was limited to 15 feet per second, an image rate of 125 frames per second was used.

Results

In steady state flow fields, downwash curves are generally symmetrical about the origin. There is upwash at negative angles of attack and downwash at positive angles of attack as shown in Fig. 2; however, in unsteady flow fields the downwash curves are shifted to the right with variable magnitudes of upwash and downwash. Because of this shift, as can be seen in Fig. 3, a significant amount of upwash can be produced on the aft wing at positive angles of attack. In Ref. 2, Burkhalter states "... a forward wing producing lift can produce upwash on an aft located wing increasing its potential to produce lift." In order to physically see the vortex interaction which caused this upwash, discrete planes in the flow field were illuminated. After several initial experiments, it became apparent that a single illuminated plane could yield significant results; therefore, the two-thirds span position was selected for reasons of visual clarity. Two oscillatory frequencies, 0.2 and 0.6 Hertz, were utilized in the tests with the 0.6 Hertz case producing the better results. For this reason, an oscillatory frequency of 0.6 Hertz was used in all cases presented in this report.

The flow fields of three different cases were analyzed in this experiment. The forward wing was set at 0, +10 and -10 degree incidences, respectively, while maintaining a zero degree incidence on the aft wing (Fig. 4). Note that Fig. 4 corresponds to the system setup in all photographs. Increasing angles of attack are considered positive in the counter clockwise direction, and positive normal forces are in the downward direction. The angle

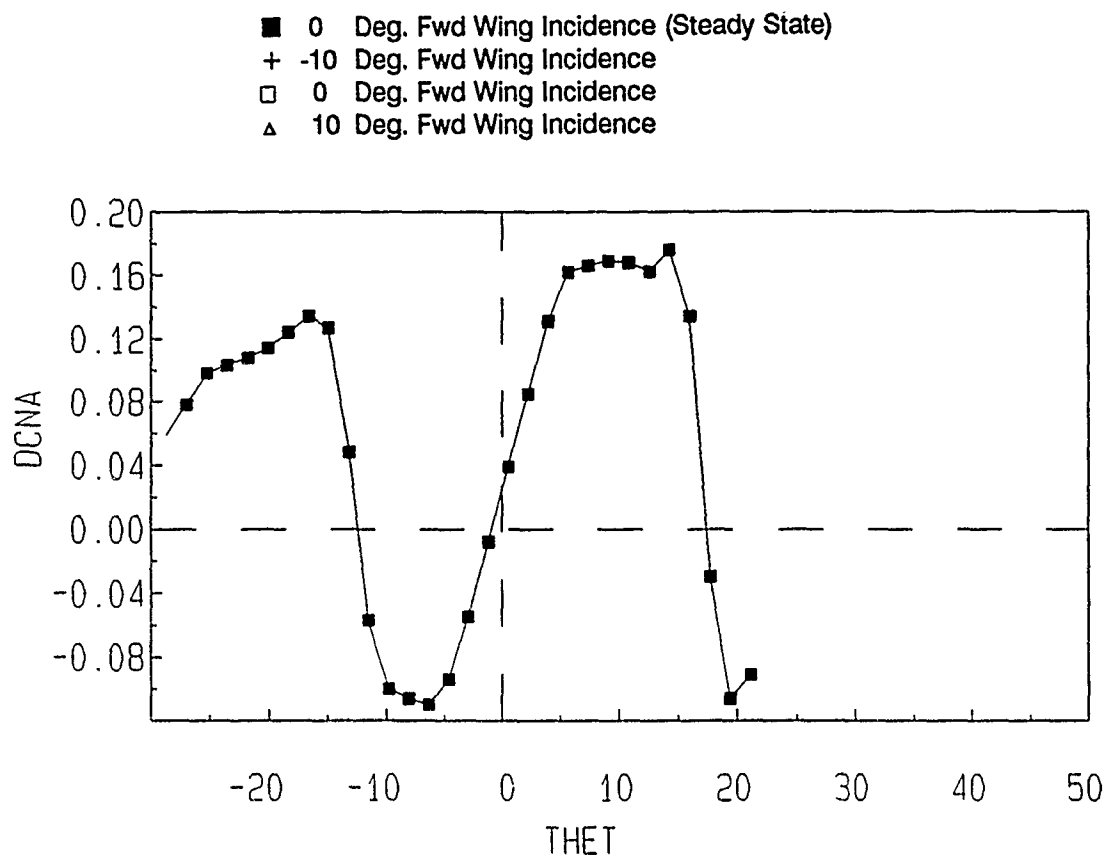


Figure 2. Steady State Downwash Coefficient versus Angle of Attack (Ref. 2)

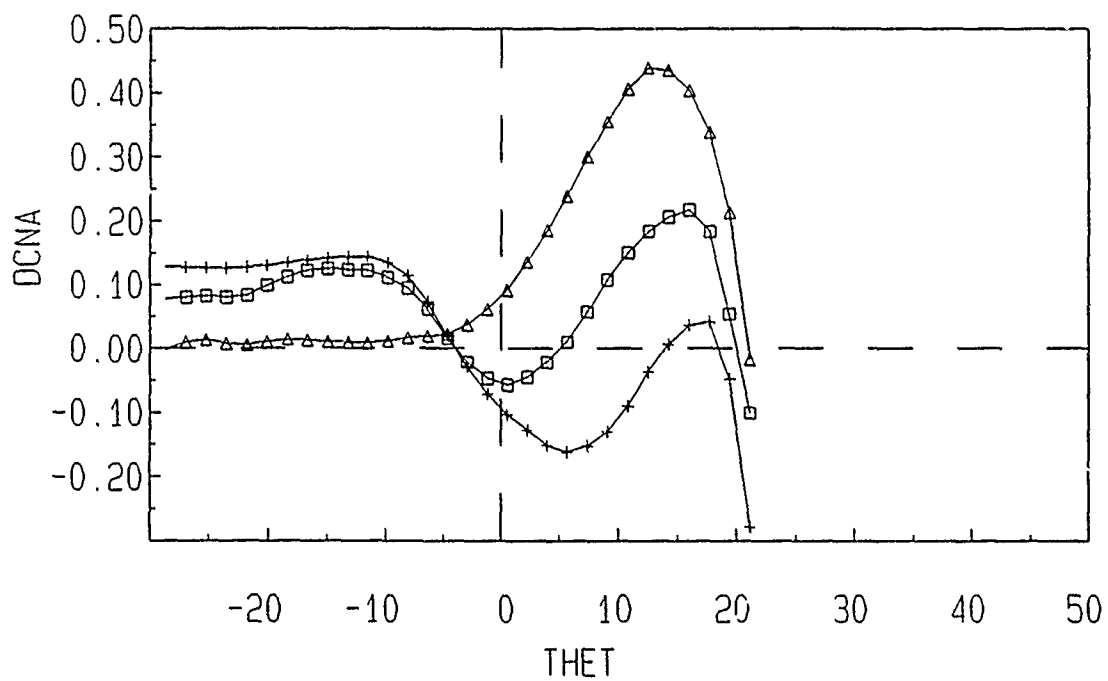


Figure 3. Unsteady Downwash Coefficient versus Angle of Attack (Ref. 2)

of attack of the system was allowed to oscillate from -35.4 degrees to $+24.6$ degrees. The area of interest in the present tests was focused on how the change in forward wing incidence affected vortex convection and interaction; as well as, how this affected the downwash on the aft wing. The unsteady downwash curve (Fig. 3) shows diverse results depending on the incidence of the forward wing. In order to better understand the mechanisms which caused these results, a correlation between the quantitative data of Ref. 2 and the qualitative flow visualization data of this report was made for the cases as discussed below.

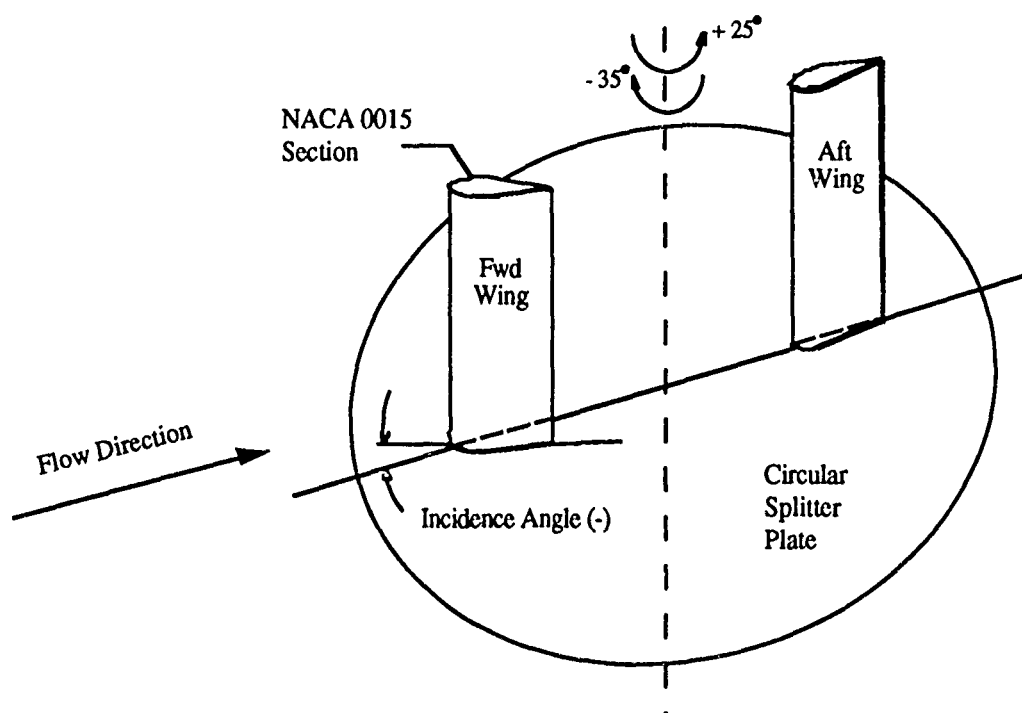


Figure 4. Schematic of System Setup

Case 1

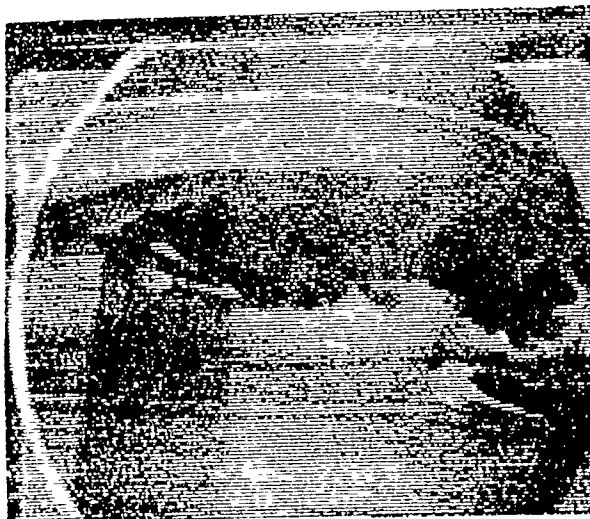
The first case that was analyzed was with the forward wing incidence set to zero degrees. With this configuration, leading edge and trailing edge vortices are seen to develop on the forward wing, which propagate downstream and interact with the aft wing. This interaction of the vortices with the aft wing is primarily dependent on the angle of



-30 degrees



-5 degrees



-20 degrees



0 degrees



-10 degrees



+5 degrees

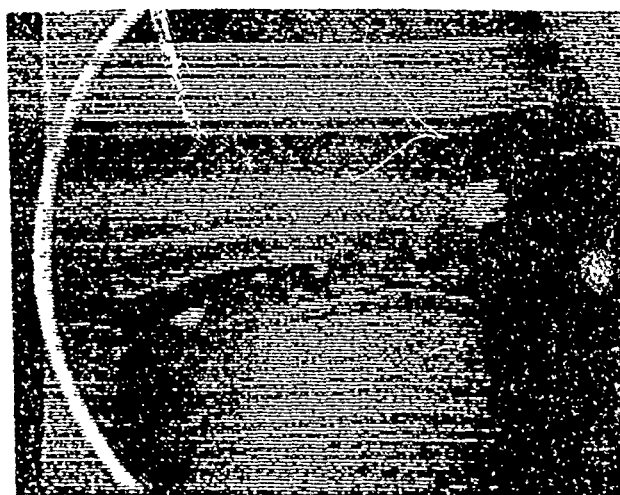
Figure 5. History of Vortex Interaction with Forward Wing Set to Zero Degrees Incidence



+10 degrees



+15 degrees



+20 degrees

Figure 6. History of Vortex Interaction with Forward Wing Set to Zero Degrees Incidence

attack of the system as well as the incidence angle of the forward wing. At higher negative angles of attack the trailing edge vortex is seen to be the dominate factor controlling the downwash on the aft wing. Fig. 5 shows a typical pitch up motion of the canard/wing system. As can be seen in these photographs, at an angle of attack of -30 degrees, the leading edge vortex does not contribute significantly to the downwash on the aft wing; however, the trailing edge vortex does seem to convect toward the aft wing suggesting that this vortex may be producing a reduced pressure on the lower surface of the aft airfoil. Because of the counter-clockwise rotation of the trailing edge vortex, the speed of the flow is accelerated near the airfoil surface. This accelerated flow causes a decrease in pressure which produces a downwash on the aft wing until the system reaches an angle of attack of approximately -8 degrees.

As the angle of attack is increased, the leading edge vortex begins to convect toward the trailing edge vortex. Since the leading edge vortex circulates in the opposite direction as the trailing edge vortex, this interaction tends to cancel each other out. At an angle of attack of approximately -5 degrees, the leading edge vortex overcomes the counter-clockwise rotation of the trailing edge vortex yielding a clockwise circulation on the aft wing upper surface. This clockwise circulation decelerates the flow over the aft wing causing a higher pressure, thus producing an upwash (Fig. 5). The aft wing continues to experience upwash from the forward wing until approximately + 5 degrees. Notice from Fig. 2 that upwash does not occur at positive angles of attack under steady state conditions. This is a very significant property of the unsteady flow field.

The strength of these vortices diminishes greatly at higher angles of attack. Fig. 6 shows a small non-vortical wake behind the forward wing. This wake seems to impact the aft airfoil causing a downwash once again.

After a careful examination of the photographs in this case, it is apparent that the vortex interaction at negative angles of attack are very different from the corresponding positive angles of attack. This is due to the circulation lag mentioned earlier. In Fig. 5 at

zero degrees angle of attack, the circulation lag is readily apparent. The leading edge vortex is still producing a clockwise rotation on the upper surface of the aft wing. This circulation produces an upwash that in steady state flows would not exist at this angle of attack. The photograph denoting an angle of attack of -30 degrees shows the configuration near the beginning of the pitch up cycle. At this relatively stationary point in the cycle, the lifting surfaces tend to emulate steady state conditions; however, there is still some residual circulation lag due to past motion history. As the configuration pitches up, the vorticity around the wings continues to increase due to the wing oscillatory motion, until the vortex is shed into the free stream. Once the vortex is shed, the process begins again and circulation lag continues to develop throughout the cycle.

When the flow fields of -20 degree and +20 degree (Figs. 5 and 6) are compared, two very different flows fields are encountered. As can be seen in the photograph, at +20 degrees, stall is delayed past the static limit. This is due to the fact that the oscillation started at -35.4 degrees angle of attack. Because of this relatively long motion history, circulation lag has developed enough at this point to delay stall. This delay in stall is yet another aspect of the unsteady flow field which could be utilized in the future to extend the maneuverability of fighter aircraft.

Case 2

The second case that was analyzed was with the forward wing set to +10 degrees incidence. This configuration proved to be the least effective in producing upwash. This was evident in both the quantitative (Fig. 3) and the qualitative (Fig. 7) data. When a comparison between this case and the previous case at an angle of attack of -30 degrees is made, there is an obvious difference in the size of the leading and trailing edge vortices. Vortex interaction also is seen to be greatly diminished. For the most part, the ability of the forward wing to produce upwash on the aft wing seems to rely in large on the incidence angle of the forward wing.



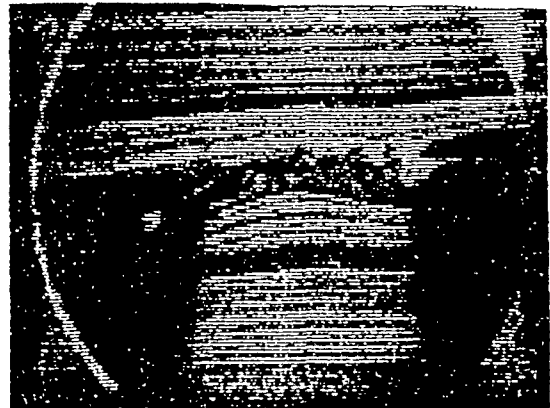
-30 degrees



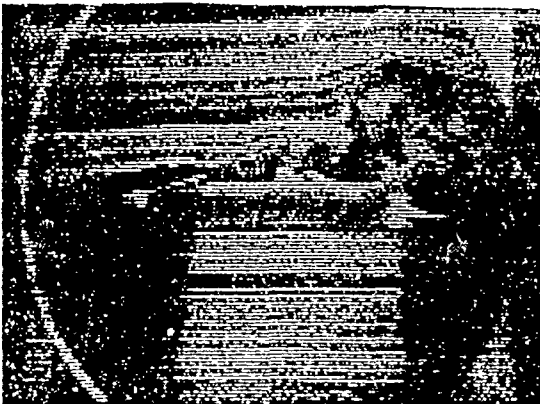
0 degrees



-15 degrees



+10 degrees



-5 degrees



+20 degrees

Figure 7. History of Vortex Interaction with Forward Wing Set to +10 Degrees Incidence

Size and interaction of these vortices are also closely dependent on the incidence angle of the forward wing. With a +10 degree incidence, the vortices created by the forward wing at the beginning of the pitch up cycle propagate downstream at a greater vertical distance from the aft wing than in the previous case. Because of the positive incidence of the forward wing, the vortices are "guided" beyond the area that can significantly influence the aft wing. This could account for the negligible downwash detected on the aft wing at high negative angles of attack (Fig. 3).

As the flow field history of this case is examined, it becomes apparent that the trailing edge vortex of the forward wing does not interact with the aft wing until approximately -5 degrees angle of attack. For this reason, neither upwash nor downwash is experienced by the aft wing in the region from -30 to -5 degrees angle of attack. The quantitative data in Fig. 3 verifies this observation by showing the downwash to be essentially zero in this region. For the remainder of the pitch up cycle, the forward wing produces only downwash on the aft wing. Although the magnitude of the downwash is not evident by the photographs, Fig. 3 indicates that this configuration produces more downwash than any of the other two cases.

At an angle of attack of -5 degrees, a slight counter clockwise circulation is produced by the trailing edge vortex. This circulation decreases the pressure on the upper surface of the aft wing producing a force in the negative direction causing a downwash. As the system continues to pitch up, the vortex interaction becomes less and less definable. At zero degrees, the forward wing produces what amounts to a non-vortical wake. No definite circulation patterns are seen in this wake; however, the wake does seem to be turbulent. Since the wake is turbulent, the velocity within the wake is lower. This creates a higher pressure on the aft wing.

The magnitude of the downwash is seen to be dependent on the trajectory angle of the non-vortical wake. The wake seems to always lie at an angle between the incidence angle of the forward wing and the angle of attack of the system. The cause of the increased

downwash could be that as the system pitches up, the trajectory angle of the wake increases as well as the wake wetted surface area of the lower surface of the aft wing. This increase in wetted area of the aft wing could cause the increase in downwash. The aft wing continues to experience this downwash until +20 degrees angle of attack.

Case 3

The third and most interesting case was with the forward wing set to -10 degrees incidence. This configuration was by far the most effective in producing upwash. The negative incidence of the forward wing seems to generate bigger and probably stronger leading and trailing edge vortices.

At the start of the pitch up cycle (-30 degree angle of attack), the counter clockwise circulation near the aft wing is controlled by the trailing edge vortex (Figure 8). As in the previous cases this causes a downwash on the aft wing. The leading edge vortex does not propagate close enough to the aft wing to affect the downwash characteristics at this point in the cycle. As the canard/wing system continues to pitch up, the leading edge vortex convects closer to the trailing edge vortex and the aft wing, thus affecting the downwash. The clockwise rotation of the leading edge vortex gradually begins to take 'control' of the trailing edge vortex. At an angle of attack of -5 degrees, the clockwise circulation from the leading edge vortex causes a higher pressure to develop and produces an upwash on the aft wing. The aft wing continues to experience upwash until approximately +15 degrees.

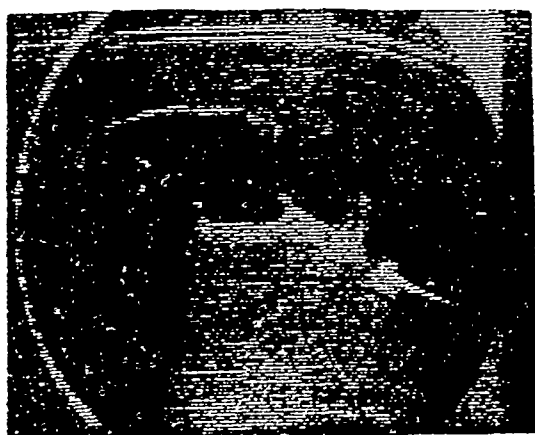
Once again, at +10 degrees the non-vortical wake forms. As before, this wake remains throughout the rest of the cycle. At this angle of attack the forward wing is essentially parallel with the free stream flow. Upwash at +10 degrees would not be possible were it not for the circulation lag. The aft wing would be experiencing pure downwash at this point in steady state conditions (no circulation lag). As the canard/wing configuration oscillates through +15 degrees, the aft wing rotates out of the non-vortical wake of the forward wing. Because of the -10 degree incidence of the



-30 degrees



-5 degrees



-20 degrees



0 degrees

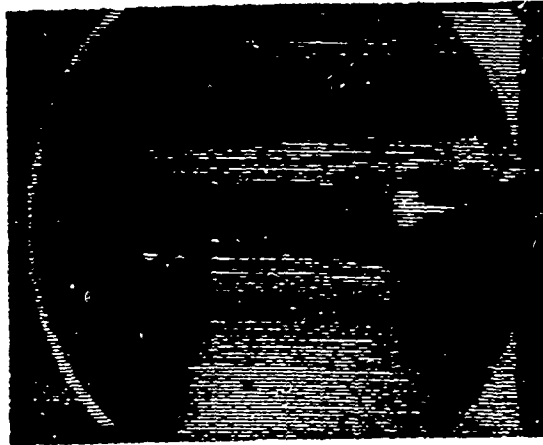


-10 degrees

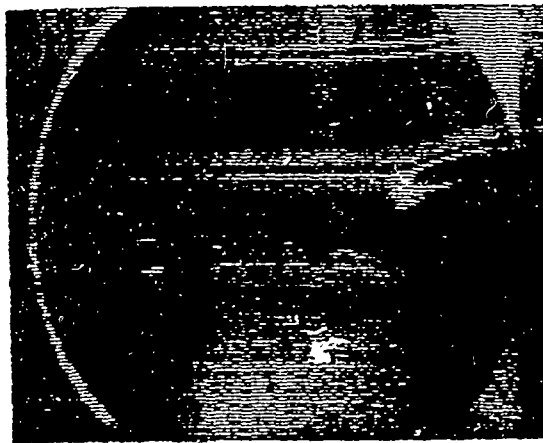


+5 degrees

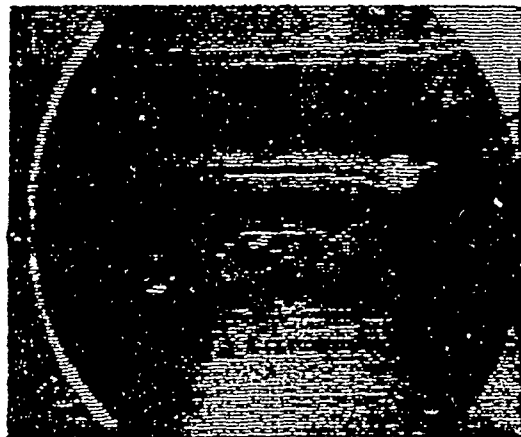
Figure 8. History of Vortex Interaction with Forward Wing Set to -10 Degrees Incidence



+10 degrees



+15 degrees



+20 degrees

Figure 9. History of Vortex Interaction with Forward Wing Set to -10 Degrees Incidence

forward wing, the wake is "guided" away from the aft wing producing less downwash than the previous cases. This characteristic of forward wing negative incidence is verified by the quantitative data in Fig. 3.

Conclusions

Unsteady flows exhibit many properties which are difficult to understand. Both quantitative and qualitative experiments have been conducted in an effort to comprehend these properties. The goal of this report was to correlate existing quantitative data with flow visualization data. Some success in identifying the flow mechanisms responsible for downwash and upwash on aft lifting surfaces was made; however, the role of circulation lag and other results are still inconclusive. The following conclusions were made from this analysis:

- (1) It is apparent from the flow visualization data that upwash on an aft lifting surface can be produced on an aft lifting surface at positive angles of attack for this configuration.
- (2) The incidence angle of the forward wing is a major contributor in producing upwash or downwash on an aft lifting surface.
- (3) Leading and trailing edge vortices are produced by the forward wing in unsteady motion. Downwash characteristics on the aft wing are closely dependent on the interaction with these vortices.
- (4) A non-vortical wake was seen to develop from the forward wing at different points in the pitch up cycle depending upon the incidence of the forward wing. Once this turbulent wake developed, no vortices were formed for the remainder of the pitch up cycle. The high pressure wake was seen to be the only flow mechanism controlling downwash or upwash beyond +10 degrees angle of attack in any case. The cause of this may be due to the circulation lag and past motion history of the system.

References

1. McAlister, Kenneth W.; Carr, Lawrence W.; and McCroskey, William J., "Dynamic Stall Experiments on the NACA 0012 Airfoil", NASA Technical Paper 1100, Ames Research Center, Moffett Field, CA., 1978.
2. Burkhalter, John E., "Downwash Measurements On A Pitching Canard-Wing Configuration", FJSRL TR - 91 - 0001, USAFA, CO., Sept. 1991.
3. Albertson, J. A., Troutt, T. R, Siuru, W. D., and Walker, J. M., "Dynamic Stall Vortex Development and the Surface Pressure Field of a Pitching Airfoil," AIAA Paper 87-1333, AIAA 19th Fluid Dynamics, Plasma Dynamics and Lasers Conference, Honolulu, Hawaii, 1987.

AN AB INITIO STUDY OF THE ADDUCTS OF ALUMINUM HYDRIDES, HALIDES, HYDROXIDES, AND OXIDES WITH HF AND HCL

Mr. Marty Wilson

ABSTRACT

The adducts of HF and HCl with aluminum hydrides, aluminum halides, aluminum hydrohalides, aluminum hydroxides, and aluminum oxides have been studied using ab initio molecular orbital methods. When the electron rich region of HF or HCl is placed between 2.0 and 3.5 angstroms above the aluminum atom, Lewis Acid/Base adducts are generally found. In a few cases, the starting geometry rearranged to give a hydrogen bonded structure. In fewer cases, chemical reactions were observed. The structures observed are reported and briefly discussed.

INTRODUCTION

Molecular Orbital Theory has been able to describe chemical bonding for several decades, but only for the past two decades, with the advent of semi-empirical and ab initio methods, have we been able to address problems of any size. Until recently, however, the size of the system which could be studied was extremely limited by computer power. With the rise of supercomputers we are now able to study large systems computationally, albeit at a primitive level if the system is very large. Supercomputers also have allowed the study of smaller systems in more sophisticated ways giving us data which can be realistically compared with experiment.

Typically, in a certain chemical environment where a great number of molecules are present, either of the same species or a different species, weak

interactions between molecules will occur. One such interaction can be classified as van der Waal's type, in which no electrons are transferred. Hydrogen bonds are a good example of this type of interaction. This interaction basically stems from London forces (induced dipole-induced dipole) or from dipole-dipole interactions. Another type of weak interaction which may be seen is the formation of Lewis Acid/Base adducts. In this case an electron rich region of one molecule donates a small fraction of its electrons to an electron poor region of another. These interactions can be intermolecular (between two different molecules) or intramolecular (between identical molecules). Lewis Acid/Base interactions, especially those which exceed thermal energy (0.9 kcal/mole at room temperature), are important because they can determine the chemical characteristics of the system as a whole. Without taking into account the possibility of these interactions, we cannot fully understand the chemistry of these systems.

Since allane is important as a high energy fuel it is of interest to the Air Force. Also, aluminum oxides constitute an important product in the exhaust of the solid state rocket boosters used for the Space Shuttle. With the recent determination that HCl is also largely present in this exhaust¹ a study of the adducts of aluminum compounds with HCl seems chemically and environmentally relevant. Since HF is chemically similar to HCl and is computationally more convenient, adducts with it also seem to be worth studying. Though we did not appreciate this upon first arriving at the F. J. Seiler Laboratory, there is considerable interest in the application of aluminum halide salts as solvents for high energy electrochemical cells. These salts are eutectic mixes prepared by mixing AlCl_3 and a Cl^- containing salt, such as NaCl or RCl, where R is an organic quaternary ammonium salt. The reaction between AlCl_3 and Cl^- is a Lewis

Acid/Base reaction such as considered in this study. Thus, the proposed study of the HF and HCl adducts of aluminum hydrides, halides, hydroxides, and oxides is of some interest to the Air Force. Add to this my research advisor's (Dr. Gilbert J. Mains) closely related work with boron adducts², and we feel this study has far reaching implications.

DISCUSSION

Two basic molecular orbital methods are available for studying the electronic properties of molecules. One of these is the semi-empirical method in which the two electron integrals are approximated, or parameterized, using experimental data. The other one is the ab initio method in which nothing is parameterized. At the F. J. Seiler Laboratory the computers available for these calculations are a cluster of micro-Vaxes. These computers can be applied to ab initio calculations and are equipped to do so, but are much better suited for semi-empirical calculations such as those performed in MOPAC³ which was developed at Seiler by Dr. J. J. P. Stewart.

All of my calculations were performed using ab initio methods, although Dr. Mains experimented with MOPAC and was sufficiently impressed with its speed and versatility so that we plan to adopt it for all future preliminary screening. The ab initio calculations were performed using the GAUSSIAN 86⁵ package developed by Pople et al. The adducts were initially explored using the 3-21G* basis set to determine approximate equilibrium structures and these were refined using the 6-31G* basis set. In order to ensure that the stationary points were minima at the Hartree-Fock level, vibrational frequencies were calculated analytically at the 6-31G* level. Correlation effects were then approximated at the 6-31G* optimized

geometries using fourth order Moller-Plesset⁴ perturbation theory, MP4(SDTQ). Energies so determined are generally accepted to be within 1-2 kcal/mole of experimental values.

All calculations were started with the electron rich end of the acid molecule 2.0 - 3.0 angstroms above the plane of the aluminum molecule and directly above the aluminum. This is the region where Lewis Acid/Base adducts were expected to form so if one were going to form we started with the most probable geometry. In most cases the desired adduct was found. However, in some cases unexpected and interesting results were obtained. For instance, there were cases where rearrangement/reaction occurred rather than adduct formation. Due to insufficient time we were unable to explore each individual surface so we cannot claim to have found the global minimum for each interaction, but we can state that we have found true minima at the Hartree-Fock level for each case.

Upon analysis of the number of calculations required to complete this study, it becomes readily apparent that there is far too much information here to fit into one project. Therefore, the work has been divided into two projects. The first of these involves the adducts of HF and HCl with aluminum hydrides and halides, while the second focuses on the HF and HCl adducts of aluminum hydroxides and oxides. The manuscript for the first part is currently in the final stages of preparation. The second manuscript will be finished sometime early this Fall.

RESULTS

In this report we will focus mainly on the oxide and hydroxide adducts as Dr. Mains' report dealt more with the hydrides and halides. While we are in the process of studying these adducts, we have enough results in hand to give a basic overview

of the work. However, to completely discuss all of our results would be a lengthy process so we will summarize the results and refer any interested readers to the respective papers for further details.

Since both HF and HCl are dipolar and many of the molecules studied are polar, the geometry of the dipole-dipole interaction is not greatly different from that of the Lewis adducts. However, one sure indication of electron donation (adduct formation) is the distortion of the equilibrium geometry of some of the aluminum compounds. For instance, each of the linear AlXO molecules (X= H, F, or Cl; oxygen doubly bonded to the aluminum) became bent when in the presence of HF or HCl. The molecules are attempting to attain the sp^2 (planar) hybridization because of the introduction of the donated electrons.

In some cases the HF or HCl adduct rearranged to hydrogen bond with a negative region of the aluminum compound leaving the compound virtually unchanged during the optimization.

Whether adduct formation or H-bonding was observed after optimization seemed to depend on the nature of the acid (HF or HCl) and on the nature of the aluminum compound. Generally speaking, for an AlXYZ compound, when X was a hydroxyl group, the hydrogen bond product was commonly formed.

The MP4(SDTQ) dissociation energies of the adducts of HF and HCl with aluminum hydroxides and oxides have not been fully tabulated at the time of this report, but should follow closely those found for halides and hydrides, which ranged from 3.0 – 25.8 kcal/mole. For the latter case the HCl adducts were more weakly bound, which should again be the trend with the hydroxides and oxides. Again, the reader is directed to the manuscripts for further information. Alternately, the reader can contact the author at Oklahoma State University.

CONCLUSION

We conclude that the empty p orbitals in Group IIIA compounds leads to the formation of Lewis adducts between HF and HCl and the majority of the aluminum hydrides, halides, hydroxides, and oxides. Since all of the studied compounds were polar, one suspects that all of the aluminum compounds can form hydrogen bonds with HF and HCl. Based on these studies one can infer that HCl is hydrogen bonded to aluminum oxides/hydroxides in the exhaust of rocket boosters. This is a very important aspect because of the concern about the fate of HCl gas in rocket exhaust.

We feel we have barely scratched the surface of the possible adducts between aluminum compounds and HF and HCl. Each of the adducts found deserves a more detailed study and a careful exploration of the energy surface. Once this is done it is possible that, for some of the adducts, a different type of interaction will be found than was found with our particular starting geometry. Also, future work should concentrate on other potential electron donors. We plan to study these other possibilities extensively in the future. If we hope to fully understand these systems, there is much work left to be done.

REFERENCES

1. P. Brimblecombe and S. L. Clegg, *J. Propulsion*, 1990, 6, 115.
2. G. J. Mains, *J. Phys. Chem.*, 1991, 95, 5089.
3. MOPAC: A General Molecular Orbital Package, J. J. P. Stewart, Quantum Chemistry Program Exchange, No. 455, (1983), Version 6 (1990).
4. C. Moller and M. S. Plesset, *Phys. Rev.*, 1934, 46, 678.
5. GAUSSIAN 86. M. J. Frisch; J. S. Binkley; H. B. Schlegel; K. Raghavachari; R. Melius; R. L. Martin; J. J. P. Stewart; C. M. Rohlfing; L. R. Kahn; D. J. Defrees; R. Seeger; R. A. Whiteside; D. J. Fox; E. M. Fleuder; J. A. Pople. Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburg, PA 15213, 1984.